

Mon Apr 3 08:24:16 2000

us-08-852-495c-1_copy_1_30000.rge

Page 2

```
KEYWORDS histone 1; promoter.
SOURCE Homo sapiens DNA.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 85)
AUTHORS Gallinari, P., La Bella, F. and Heintz, N.
TITLE Characterization and purification of H1np2, a novel CCAAT-binding
protein that interacts with a histone H1 subtype-specific consensus
element.
JOURNAL Mol. Cell. Biol. 9 (4), 1566-1575 (1989)
MEDLINE 89261784
COMMENT On Jul 23, 1994 this sequence version replaced gi:341438.
FEATURES
SOURCE
1. 85
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Hela"
/protein_bind
28..39
/bound_moiety="SP1"
CAAT_signal
53..70
/gene="HIS1"
53..85
/gene="HIS1"
77..85
TATA_signal
/gene="HIS1"
BASE COUNT 26 a 27 c 18 g 14 t
ORIGIN
Query Match 0.2%; Score 46; DB 10; Length 85;
Best Local Similarity 100.0%; Pred. No. 9,6e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26024 GTGCTAGGTGATGACCAATCAACAGCGCCCTACCTATATTAAG 26069
|||||
DB 40 GTGCTAGGTGATGACCAATCAACAGCGCCCTACCTATATTAAG 85
|||||

RESULT 2
AF042300 73 bp DNA STS 26-JAN-1999
LOCUS Homo sapiens chromosome 6 map 6p21.3, sequence tagged site.
DEFINITION AF042300
ACCESSION AF042300.1 GI:4191287
VERSION STS.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 73)
AUTHORS Lin, L., Jin, L., Kimura, A., Carrington, M. and Mignot, E.
TITLE DQ microsatellite association studies in three ethnic groups
JOURNAL Tissue Antigens 50 (5), 507-520 (1997)
MEDLINE 96049194
REFERENCE 2 (bases 1 to 73)
AUTHORS Lin, L., Jin, L., Lin, X., Voros, A., Underhill, P. and Mignot, E.
TITLE Microsatellite single nucleotide polymorphisms in the HLA-DQ region
JOURNAL Tissue Antigens 52 (1), 9-18 (1998)
REFERENCE 3 (bases 1 to 73)
AUTHORS Lin, L., Jin, L., Lin, X., Voros, A., Underhill, P. and Mignot, E.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-1998) Psychiatry and Behavioral Sciences,
Stanford University, Lab Surge Building, 1201 Welch Road, Room
P126, Palo Alto, CA 94304, USA
FEATURES
SOURCE
1. 73
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
repeat_region
1. 73
repeat_region
1. 73
```

```
/note="microsatellite"
/rpt_type="tandem"
BASE COUNT 32 a 29 c 3 g 9 t
ORIGIN
Query Match 0.1%; Score 40; DB 13; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACACACACACACACACACACACACACACACACAAATG 40
|||||
DB 19 CACACACACACACACACACACACACACACACACAAATG 58
|||||

RESULT 4
AF042302 77 bp DNA STS 26-JAN-1999
LOCUS Homo sapiens chromosome 6 map 6p21.3, sequence tagged site.
DEFINITION AF042302
ACCESSION AF042302.1 GI:4191289
VERSION STS.
KEYWORDS
```

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS Lin, L., Jin, L., Kimura, A., Carrington, M., and Mignot, E.
TITLE DO microsatellite association studies in three ethnic groups
JOURNAL Tissue Antigens 50 (5), 507-520 (1997)
MEDLINE 96049194
REFERENCE 2 (bases 1 to 77)
AUTHORS Lin, L., Jin, L., Lin, X., Voros, A., Underhill, P., and Mignot, E.
TITLE Microsatellite single nucleotide polymorphisms in the HLA-DQ region
JOURNAL Tissue Antigens 52 (1), 9-18 (1998)
MEDLINE 98378271
REFERENCE 3 (bases 1 to 77)
AUTHORS Lin, L., Jin, L., Lin, X., Voros, A., Underhill, P., and Mignot, E.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-1998) Psychiatry and Behavioral Sciences, Stanford University, Lab Surge Building, 1201 Welch Road, Room P126, Palo Alto, CA 94304, USA
FEATURES
source
1. .77
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
repeat_region
1. .77
/note="microsatellite"
/rpt_unit="ca
/rpt_type="tandem
BASE COUNT 34 a 31 c 3 g 9 t
ORIGIN
Query Match 0.18; Score 40; DB 13; Length 77;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACACACACACACACACACACACACACACACACAAATG 40
Db 21 CACACACACACACACACACACACACACACACACAAATG 60
RESULT 5
HUMUT6154A 77 bp DNA STS 28-DEC-1994
LOCUS Human STS U6154, 5' primer bind, sequence tagged site.
DEFINITION L30742
ACCESSION L30742
VERSION L30742.1 GI:604798
KEYWORDS STS; PCR primer; STS sequence; microsatellite DNA; microsatellite marker; sequence tagged site; tetranucleotide repeat.
SOURCE Homo sapiens DNA.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS 1 (bases 1 to 77)
Gerken, S.C., Matsunami, N., Plaetke, R., Albertsen, H., Ballard, L., Melis, R., Lawrence, E., Moore, M., Holik, P.R., Carlson, M., Zhao, X., Robertson, M., Bradley, P., Elsnier, T., Tingey, A., Lalouel, J.-M., and White, R.
TITLE Genetic and physical mapping of simple sequence repeat containing sequence tagged sites from the human genome
JOURNAL Unpublished (1994)
COMMENT Submitted by: Utah Center for Human Genome Research University of Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: stsecorona.med.utah.edu
Primer A: GATCGAGACACCTGGA
Primer B: TGGTATTTGAAGGAATGCC
End to Label: Primer A
PCR Profile:
Initial Denaturation: 94C 300sec

Cycles Denaturation Annealing Extension 5 94
C 10 sec. 58 C 10 sec. 72 C 20 sec. 30
54 C 10 sec. 72 C 20 sec. Mg++: 1.00 mM
Gel: Acrylamide 7%, Formamide 3%, Urea 348
Alleles: 0.
FEATURES
source
1. .77
/organism="Homo sapiens"
/db_xref="taxon:9606"
primer_bind
12. .29
/evidence="experimental"
BASE COUNT 36 a 17 c 11 g 11 t 2 others
ORIGIN
Query Match 0.18; Score 40; DB 13; Length 77;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24485 CACATGATGAAACCATCTCTCTAAATACAAAAA 24524
Db 31 CACATGATGAAACCATCTCTCTAAATACAAAAA 70
RESULT 6
MMD4NDS1 97 bp DNA ROD 25-JUN-1992
LOCUS M.musculus Dndsl microsatellite DNA.
DEFINITION X55208
ACCESSION X55208
VERSION X55208.1 GI:50662
KEYWORDS microsatellite DNA.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 97)
Hearne, C.M.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1990) Hearne C.M., University of Oxford, Nuffield Dept. Surgery, Level 6, John Radcliffe Hosp., Headington, Oxford, UK
REFERENCE 2 (bases 1 to 97)
Cornall, R.J., Altman, T.J., Hearne, C.M., and Todd, J.A.
TITLE The generation of a library of PCR-analyzed microsatellite variants for genetic mapping of the mouse genome
JOURNAL Genomics 10 (4), 874-881 (1991)
MEDLINE 92009923
FEATURES
source
1. .97
/organism="Mus musculus"
/strain="NOD"
/db_xref="taxon:10090"
/chromosome="4"
/map="45 CM"
BASE COUNT 41 a 27 c 6 g 23 t
ORIGIN
Query Match 0.18; Score 39; DB 12; Length 97;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACACACACACACACACACACACACACACACACAAAT 39
Db 36 CACACACACACACACACACACACACACACACACAAAT 74
RESULT 7
AV025275/c 97 bp DNA STS 02-MAR-1999
LOCUS AV025275
DEFINITION Rattus norvegicus, OTSUKA clone, OT48.13/788f08, microsatellite sequence, sequence tagged site.
ACCESSION AV025275

Mon Apr 3 08:24:16 2000

us-08-852-495c-1_copy_1_30000.rge

Page 4

VERSION	AA025275.1	GI:4515198
KEYWORDS	STS.	
SOURCE	Rattus norvegicus DNA, clone:OT48.13/788f08.	
ORGANISM	Rattus norvegicus	
REFERENCE	Eukaryota: Metazoa; Chordata: Craniata; Vertebrata: Mammalia: Eutheria: Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
AUTHORS	1 (sites)	
TITLE	Watanabe,T.K., Hishiyaki,H., Kanemoto,N., Miyakita-Mizoguchi,A., Oga,K., Okuno,S., Ono,T., Tsuji,A., Hayashi,H., Adachi,M., Yamashita,Y., Iriye,Y., Takahashi,E., Takagi,T., Nakamura,Y. and Iamaguchi,A.	
JOURNAL	The large-scale mapping of rat microsatellite markers	
REFERENCE	Unpublished (1998)	
AUTHORS	2 (bases 1 to 97)	
TITLE	Watanabe,T.K.	
JOURNAL	Direct Submmission	
FEATURES	Submitted (24-JUL-1998) to the DDBJ/EMBL/GenBank databases. Takeshi K Watanabe, Otsuka Pharmaceutical Co., Ltd., Otsuka GEN Research Institute, 463-10, Kagasuno, Kawachi-cho, Tokushima, Tokushima 771-0192, Japan (E-mail:watanabe@otsuka.genome.ad.jp, Tel:81-886-65-2888, Fax:81-886-37-1035)	
SOURCE	Location/Qualifiers	
	1..97	
	/organism="Rattus norvegicus"	
	/db_xref="taxon:10116"	
	/clone="OT48.13/788f08"	
	/note="OT48.13/788f08F-5'-TCCAGGCCCTAGATTACATTTA-3', OT48.13/788f08R-5'-CCGCGAGCATTAATTAGGC-3'"	
BASE COUNT	16 a 12 c 50 g 39 t	
ORIGIN		

Query Match	0.1%;	Score 39;	DB 13;	Length 97;
Best Local Similarity	100.0%;	Pred. No. 3.5e+08;		
Matches	39;	Conservative	0;	Mismatches 0; Indels 0; Gaps
Oy	1	CACACACACACACACACACACACACACACACACAAAT	39	
Db	62	CACACACACACACACACACACACACACACACACAAAT	24	
RESULT	8			
LOCUS	131337			
DEFINITION	Sequence 249 from patent US 5582979.	PAT	06-FEB-1997	
ACCESSION	131337			
VERSION	131337.1	GI:1822128		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 44)			
TITLE	Weber, J.L.			
JOURNAL	Length polymorphisms in (dc-da).sub.n.(dc-dr).sub.n sequences and			
FEATURES	method of using the same			
	Patent: US 5582979 A 249 10-DEC-1996;			
	Location/Qualifiers			
	1..44			
	/organism="unknown"			
BASE COUNT	23 a 21 c 0 g 0 t			
ORIGIN				

	0.1%;	Score 38;	DB 5;	Length 44;
Query Match	Similarity 100.0%;	Pred. No. 1.2e-07;		
Best Local	Matches 38;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0.
OY	1 CACACACACACACACACACACACACACAAAAA	38		
Db	2 CACACACACACACACACACACACACACAAAAA	39		

RESULT 9

[illegible]

	RESULT	10
	MMDND8/c	
	LOCUS	
	DEFINITION	
	ACCESSION	
	VERSION	
	KEYWORDS	
	SOURCE	
	ORGANISM	
	REFERENCE	
	AUTHORS	
	TITLE	
JOURNAL	JOURNAL	
MEDLINE	JOURNAL	
REFERENCE	REFERENCE	
AUTHORS	AUTHORS	
TITLE	TITLE	
JOURNAL	JOURNAL	
DEPT.	DEPT.	
FEATURES	source	
BASE COUNT	satellite	
ORIGIN	satellite	

MMDND8
M.musculus microsatellite DNA Dond8.
X55926.1 GI:50733
microsatellite DNA.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 87)
Cornall,R.J., Altman,T.J., Hearne,C.M. and Todd,J.A.
The generation of a library of PCR-analyzed microsatellite variants
for genetic mapping of the mouse genome
Genomics 10 (4), 874-881 (1991)

2 (bases 1 to 87)
Hearne,C.M.
Direct Submission
Submitted (22-OCT-1990) C.M. Hearne, UNIVERSITY OF OXFORD, NUFFIELD
DEPT. SURGERY, LEVEL 6, JOHN RADCLIFFE HOSP., HEADINGTON OXFORD, UK
Location/Organism
1. .87
/location="Mus musculus"
/organism="NOD,ssp.domesticus"
/db_xref="taxon:10090"
1. .87
/note="Dond8"

19 A 8 C 29 G 31 T

	Query Match	0.1%	Score 38;	DB 12;	Length 87;
	Best Local Similarity	100.0%;	Pred. No.	1.1e-07;	
Matches	38;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
07	1 CACACACACACACACACACACACACACAA	38			
Db	62 CACACACACACACACACACACACACAAA	25			

RESULT 11

HS2492G9
LOCUS HS2492G9 100 bp DNA STS 28-NOV-1994
DEFINITION H. sapiens (HS252) DNA segment containing (CA) repeat; clone
AFM492G9; single read, sequence tagged site.
ACCESSION 217138
VERSION 217138.1 GI:23763
KEYWORDS STS; CA repeat; dinucleotide repeat; GT repeat; microsatellite DNA;
microsatellite marker; repeat polymorphism.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 100)
AUTHORS Weissbach, J.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-1992) Genethon, B.P. 60, 91002 Evry Cedex France.
MEDLINE E-mail: Jean.Weissbach@genethon.fr
REFERENCE 2 (bases 1 to 100)
AUTHORS Weissbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J.,
Millesseu, P., Vaysseix, G. and Lathrop, M.
TITLE A second-generation linkage map of the human genome
JOURNAL Nature 359 (6398), 794-801 (1992)
MEDLINE 93063290
REFERENCE 3 (bases 1 to 100)
AUTHORS Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C.,
Millesseu, P., Marc, S., Bernardi, G., Lathrop, M. and Weissbach, J.
TITLE The 1993-94 Genethon human genetic linkage map
JOURNAL Nat. Genet. 7 (2 Spec No), 246-339 (1994)
MEDLINE 95004593
COMMENT cloning vector is M13mp18;
full automatic.
FEATURES
source 1.100
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/cell_line="CEPH 134702"
/clone_1lb="genomic DNA"
BASE COUNT 35 a 31 c 10 g 24 t
ORIGIN
Query Match 0.1%; Score 38; DB 13; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CACACACACACACACACACACACACACACAA 38
Db 28 CACACACACACACACACACACACACACACAA 65
RESULT 12
MMVIMV11 102 bp DNA ROD 30-AUG-1996
LOCUS MMVIMV11 102 bp DNA ROD 30-AUG-1996
DEFINITION M.musculus DNA for vimentin-binding fragment VII.
ACCESSION X89128
VERSION X89128.1 GI:872135
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 102)
AUTHORS Wang, X., Tolstosov, G., Shoeman, R.L. and Traub, P.
TITLE Selective binding of specific mouse genomic DNA fragments by mouse
vimentin filaments in vitro
JOURNAL DNA Cell Biol. 15 (3), 209-225 (1996)
MEDLINE 96226400
REFERENCE 2 (bases 1 to 102)
AUTHORS Shoeman, R.L.
TITLE Direct Submission
JOURNAL Submitted (23-JUN-1995) Shoeman R. L., Max-Planck-Institute fuer
Zellbiologie, Rosenhof, Ladenburg, Germany, D-66526

FEATURES
source 1.102
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="Ehrlich ascites tumor cell"
protein_bind 1.102
/standard_name="fragment VII in citation #1"
/citation=[1]
/bound_molecule="vimentin"
/function="sequence bound by vimentin filaments in vitro"
evidence=experimental
repeat_region 14..55
/citation=[1]
/rpt_family="d(CA/GT)n dinucleotide repeats"
/rpt_unit=14..15
/function="recombination, potentially to form non-B DNA"
/rpt_type=DIRECT
BASE COUNT 16 a 17 c 32 g 37 t
ORIGIN
Query Match 0.1%; Score 38; DB 12; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CACACACACACACACACACACACACACACAA 38
Db 48 CACACACACACACACACACACACACACACAA 11
RESULT 13
AU046973 104 bp DNA STS 22-APR-1999
LOCUS AU046973 104 bp DNA STS 22-APR-1999
DEFINITION Rattus norvegicus, Otsuka clone, 109c04, microsatellite sequence,
sequence tagged site.
ACCESSION AU046973
VERSION AU046973.1 GI:4631608
KEYWORDS STS.
SOURCE Rattus norvegicus (strain:Brown Norway) liver hepatocyte DNA,
clone:109c04.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (sites)
AUTHORS Watanabe, T.K., Hishiyaki, H., Okuno, S., Mizoguchi, A., Oga, K.,
Tsuji, A., Ono, T., Yamashiki, Y., Kanemoto, N., Takahashi, E., Irie, Y.,
Nakamura, Y., Takagi, Y. and Tanigami, A.
TITLE The large-scale mapping of rat microsatellite markers
JOURNAL Unpublished (1998)
REFERENCE 2 (bases 1 to 104)
AUTHORS Watanabe, T.K.
TITLE Direct Submission
JOURNAL Submitted (11-DEC-1998) to the DDBJ/EMBL/GenBank databases. Takeshi
K Watanabe, Otsuka GEN Research Institute, Otsuka Pharmaceutical
Co., Ltd, 463-10, Kagasuno, Kawasuno, Chokushi, Tokushima
771-0192, Japan (E-mail: watanabe@otsuka.gr.jp, Tel:+81-886-65-2888,
Fax:+81-886-37-1035)
FEATURES
source 1.104
Location/Qualifiers
/organism="Rattus norvegicus"
/strain="Brown Norway"
/db_xref="taxon:10116"
/cell_type="hepatocyte"
/clone="109c04"
/tissue_type="liver"
/note="109c04F-5'-CCCTTCCTATTTC-3',
109c04R-5'-ATTAAAGTATAGGAATACACAT-3'."
BASE COUNT 15 a 14 c 21 g 54 t
ORIGIN
Query Match 0.1%; Score 38; DB 13; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 ACACACACACACACACACACACACAAATGA 41
|||||
Db 73 ACACACACACACACACACACACACAAATGA 36

RESULT 14

AF042298 69 bp DNA STS 26-JAN-1999
LOCUS Homo sapiens chromosome 6 map 6p21.3, sequence tagged site.
DEFINITION AF042298
ACCESSION AF042298.1 GI:4191285
VERSION STS.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 69)
AUTHORS Lin, L., Jin, L., Kimura, A., Carrington, M. and Mignot, E.
TITLE DQ microsatellite association studies in three ethnic groups
JOURNAL Tissue Antigens 50 (5), 507-520 (1997)
MEDLINE 98049194

REFERENCE 2 (bases 1 to 69)
AUTHORS Lin, L., Jin, L., Lin, X., Voros, A., Underhill, P. and Mignot, E.
TITLE Microsatellite single nucleotide polymorphisms in the HLA-DQ region
JOURNAL Tissue Antigens 52 (1), 9-18 (1998)
MEDLINE 98378271

REFERENCE 3 (bases 1 to 69)
AUTHORS Lin, L., Jin, L., Lin, X., Voros, A., Underhill, P. and Mignot, E.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-1998) Psychiatry and Behavioral Sciences,
Stanford University, Lab Surge Building, 1201 Welch Road, Room
P126, Palo Alto, CA 94304, USA

FEATURES

Source
1..69
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
/note="microsatellite"
repeat_region
1..69
/rpt_unit="ca"
/rpt_type="tandem"

BASE COUNT 30 a 27 c 9 g 9 t
ORIGIN

Query Match

0.1%; Score 37; DB 13; Length 69;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 ACACACACACACACACACACACACAAATG 40
|||||
Db 16 ACACACACACACACACACACACACAAATG 52

RESULT 15

AF042305 83 bp DNA STS 26-JAN-1999
LOCUS Homo sapiens chromosome 6 map 6p21.3, sequence tagged site.
DEFINITION AF042305
ACCESSION AF042305.1 GI:4191292
VERSION STS.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 83)
AUTHORS Lin, L., Jin, L., Kimura, A., Carrington, M. and Mignot, E.
TITLE DQ microsatellite association studies in three ethnic groups
JOURNAL Tissue Antigens 50 (5), 507-520 (1997)
MEDLINE 98049194

REFERENCE 2 (bases 1 to 83)
AUTHORS Lin, L., Jin, L., Lin, X., Voros, A., Underhill, P. and Mignot, E.
TITLE Microsatellite single nucleotide polymorphisms in the HLA-DQ region
JOURNAL Tissue Antigens 52 (1), 9-18 (1998)
MEDLINE 98378271

REFERENCE 3 (bases 1 to 83)
AUTHORS Lin, L., Jin, L., Lin, X., Voros, A., Underhill, P. and Mignot, E.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-1998) Psychiatry and Behavioral Sciences,
Stanford University, Lab Surge Building, 1201 Welch Road, Room
P126, Palo Alto, CA 94304, USA

FEATURES

Source
1..83
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
/note="microsatellite"
repeat_region
1..83
/rpt_unit="ca"
/rpt_type="tandem"

BASE COUNT 37 a 34 c 3 g 9 t
ORIGIN

Query Match

0.1%; Score 37; DB 13; Length 83;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CACACACACACACACACACACACACACAA 37
|||||
Db 19 CACACACACACACACACACACACACACAA 55

Search completed: March 29, 2000, 10:09:37
Job time: 24326 sec

Mon Apr 3 08:24:16 2000

us-08-852-495c-1_copy_1_30000.rge

Page 7



9
1
2



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 03:50:02 ; Search time 599.56 Seconds

(without alignments)
12518.785 Million cell updates/sec

Title: US-08-852-495c-1_COPY_1_30000
Perfect score: 30000

Sequence: 1 CACACACACACACACACACA.....AGCCATACAGAGTGGCCCC 30000

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 311585 seqs, 125096042 residues

Word size: 0

Total number of hits satisfying chosen parameters: 433264

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database: N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	0.1	44	1 T65763	Repeat sequence fr
2	38	0.1	51	1 T65773	Repeat sequence fr
3	37	0.1	60	1 Q33930	Microsatellite seq
4	37	0.1	70	1 Q33849	Microsatellite seq
5	36	0.1	36	1 Q33828	Microsatellite seq
6	36	0.1	36	1 Q33882	Microsatellite seq
7	36	0.1	36	1 Q33953	Microsatellite seq
8	36	0.1	37	1 Q34041	Microsatellite seq
9	36	0.1	37	1 Q34178	Microsatellite seq
10	36	0.1	37	1 Q33900	Microsatellite seq
11	36	0.1	37	1 T65732	Repeat sequence fr
12	36	0.1	38	1 T66048	(dc-da)n.(dg-dt)n
13	36	0.1	38	1 T65750	Repeat sequence fr
14	36	0.1	39	1 Q33648	Microsatellite seq
15	36	0.1	39	1 Q33807	Microsatellite seq
16	36	0.1	39	1 Q33825	Microsatellite seq
17	36	0.1	39	1 T65731	Repeat sequence fr
18	36	0.1	40	1 Q34091	Sequence of a micr
19	36	0.1	40	1 T65725	Repeat sequence fr
20	36	0.1	40	1 T66051	(dc-da)n.(dg-dt)n
21	36	0.1	40	1 T66054	(dc-da)n.(dg-dt)n
22	36	0.1	41	1 T65745	Repeat sequence fr
23	36	0.1	41	1 T65710	Repeat sequence fr
24	36	0.1	41	1 T65758	Repeat sequence fr
25	36	0.1	42	1 Q33770	Microsatellite seq
26	36	0.1	42	1 T65797	Repeat sequence fr
27	36	0.1	42	1 T65757	Repeat sequence fr
28	36	0.1	43	1 T65794	Repeat sequence fr
29	36	0.1	44	1 Q34113	Sequence of a micr
30	36	0.1	44	1 Q33636	Microsatellite seq
31	36	0.1	44	1 Q33983	Microsatellite seq
32	36	0.1	44	1 T65749	Repeat sequence fr
33	36	0.1	44	1 T65761	Repeat sequence fr

ALIGNMENTS

C	34	36	0.1	45	1	Q33968	Microsatellite seq
C	35	36	0.1	45	1	Q33915	Microsatellite seq
C	36	35	0.1	46	1	Q33840	Microsatellite seq
C	37	36	0.1	46	1	Q33939	Microsatellite seq
C	38	36	0.1	46	1	T65719	Repeat sequence fr
C	39	36	0.1	46	1	T65709	Repeat sequence fr
C	40	36	0.1	46	1	T65756	Repeat sequence fr
C	41	36	0.1	46	1	T65780	Repeat sequence fr
C	42	36	0.1	47	1	Q33834	Microsatellite seq
C	43	36	0.1	47	1	T65716	Repeat sequence fr
C	44	36	0.1	47	1	T65713	Repeat sequence fr
C	45	36	0.1	47	1	T65755	Repeat sequence fr

Query Match 0.1%; Score 38; DB 1; Length 44;
Best Local Similarity 100.0%; Pred. No. 0.00062;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
T65773
ID T65773; standard; DNA; 51 BP.
AC 17-JUN-1997 (first entry)
DE Repeat sequence from polymorphic marker clone Mf6102.
KW Polymorphism; repeat sequence; genetic marker; primer; amplification;
KW PCR; polymerase chain reaction; paternity; maternity; human; pedigree;

KW linkage analysis; genetic disease; animal; plant; breeding; locus;
KM hybridisation; chromosome; ds.
OS US582979-A.
PN 10-DEC-1996.
PF 21-APR-1989; 341562.
PR 21-APR-1989; US-341562.
PR 05-SEP-1991; US-754351.
PR 04-APR-1994; US-222177.
PA (MARS-) MARSHFIELD CLINIC.
PI Weber JL.
DR WPI: 97-042299/04.
PR Detection of polymorphic genetic markers of the form
PT (dc-da)n(dc-dt)n - using novel nucleic acid mols. as primers
PS Claim 1; Column 13-14; 186pp; English.
CC The invention relates to the isolation of polymorphic repeat sequences
CC having the sequence (dc-da)n.(dc-dt)n which can be used as genetic
CC markers. Primers based on these sequences can be used to detect these
CC repeats. especially for use in e.g. paternity or maternity testing.
CC human genetic analysis such as linkage analysis of genetic disease,
CC commercial animal or plant breeding or pedigree analysis. clones
CC containing the repeat sequences were isolated by hybridisation of
CC chromosome-specific phage libraries with a synthetic poly(dc-da) (dc-dt)
CC probe. Over 100 repeat blocks were isolated. The inserts from the clones
CC were amplified by primers 165798-166047. These clones where the repeat
CC sequence has been determined are shown in 165704-1797. This repeat
CC sequence is from the marker clone Kd1102 which contains the repeat
SQ Sequence 51 BP; 27 A; 24 C; 0 G; 0 U;

Query Match 0.1%; Score 38; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.00061;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACACACACACACACACACACACACACACAA 38
DB 4 CACACACACACACACACACACACACACACAA 41

RESULT 3
Q33930/c
ID Q33930 standard; DNA: 60 BP.
AC Q33930:
DE 02-FEB-1993 (first entry)
PT Microsatellite sequence from clone TGA1337.
KM PCR: selection; primers; OPTIPRIM; breeding; cattle; parentage;
KM genetic mapping; traits; amplification; ss.
OS Bos taurus.
PN WO9213102-A.
PF 06-AUG-1992.
PF 15-JAN-1992; US-000340.
PA (GENM-) GENMARK.
PI Georges M. Massey JM.
DR WPI: 92-284684/34.
PR Polymorphic bovine DNA markers - used in genetic identification,
PT gene mapping, and selective breeding
PS Table 7; Page 301; 517pp; English.
CC The sequence is that of a bovine microsatellite sequence obt'd. by
CC screening a library of bovine MboI DNA fragments of between
CC 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MboI sites, the frequency of
CC (T6)n >9 microsatellites in the bovine genome is estimated at >100.
CC 000 The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for in vitro
CC amplification of the corresp. microsatellite (using the program
CC OPTIPRIM). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved in the determination of economically important

CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SQ Sequence 60 BP; 2 A; 0 C; 27 G; 31 T;

Query Match 0.1%; Score 37; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ACACACACACACACACACACACACACACAA 38
DB 41 ACACACACACACACACACACACACACACAA 5

RESULT 4
Q33849/c
ID Q33849 standard; DNA: 70 BP.
AC Q33849:
DE 02-FEB-1993 (first entry)
PT Microsatellite sequence from clone TGA26.
KM PCR: selection; primers; OPTIPRIM; breeding; cattle; parentage;
KM genetic mapping; traits; amplification; ss.
OS Bos taurus.
PN WO9213102-A.
PF 06-AUG-1992.
PF 15-JAN-1992; US-000340.
PA (GENM-) GENMARK.
PI Georges M. Massey JM.
DR WPI: 92-284684/34.
PR Polymorphic bovine DNA markers - used in genetic identification,
PT gene mapping, and selective breeding
PS Table 7; Page 269; 517pp; English.
CC The sequence is that of a bovine microsatellite sequence obt'd. by
CC screening a library of bovine MboI DNA fragments of between
CC 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MboI sites, the frequency of
CC (T6)n >9 microsatellites in the bovine genome is estimated at >100.
CC 000 The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for in vitro
CC amplification of the corresp. microsatellite (using the program
CC OPTIPRIM). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved in the determination of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SQ Sequence 70 BP; 1 A; 1 C; 33 G; 35 T;

Query Match 0.1%; Score 37; DB 1; Length 70;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACACACACACACACACACACACACACACAA 37
DB 37 CACACACACACACACACACACACACACACAA 1

RESULT 5
Q33828/c
ID Q33828 standard; DNA: 36 BP.
AC Q33828:
DE 02-FEB-1993 (first entry)
PT Microsatellite sequence from clone TGA23.
KM PCR: selection; primers; OPTIPRIM; breeding; cattle; parentage;
KM genetic mapping; traits; amplification; ss.
OS Bos taurus.
PN WO9213102-A.
PF 06-AUG-1992.
PF 15-JAN-1992; US-000340.

PR 15-JAN-1991: US-642342.
 PA (GENM-) GENMARK.
 PI Georges M, Massey JM;
 DR WPI: 92-284684/34.
 PT Polymorphic bovine DNA markers - used in genetic identification,
 PS gene mapping, and selective breeding
 Table 7: Page 261: 517pp: English.
 CC The sequence is that of a bovine microsatellite sequence obtd. by
 CC screening a library of bovine MboI DNA fragments of between
 CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
 CC One out of 50 clones cross-hybridised. Assuming independent
 CC distribution of microsatellites and MboI sites, the frequency of
 CC (T6)n > 9 microsatellites in the bovine genome is estimated at >100,
 CC 000. The sequence information for ca. 230 such bovine microsatellites
 CC is summarised in the specification and indexed herein (see below).
 CC The sequences upstream and downstream of the microsatellite sequence
 CC were used to generate the required PCR primers for in vitro
 CC amplification of the corresp. microsatellite (using the program
 CC OPTIPRIM). The microsatellites may be used to identify individuals,
 CC for parentage testing, and in the genetic mapping of economic trait
 CC loci, or genes involved in the determination of economically important
 CC traits esp. in cattle, to allow selective breeding.
 CC See also Q33501-34437.
 SQ Sequence 36 BP; 0 A; 18 G; 18 T;

Query Match 0.1%; Score 36; DB 1; Length 36;
 Best Local Similarity 100.0%; Pred. No. 0.0031;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CACACACACACACACACACACACACACA 36
 Db 36 CACACACACACACACACACACACACACA 1

RESULT 6
 ID Q33882 standard; DNA; 36 BP.
 AC Q33882;
 DT 02-FEB-1993 (first entry)
 DE Microsatellite sequence from clone TGLA303.
 KW PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;
 KM genetic mapping; traits; amplification; ss.
 OS Bos taurus.
 PN WO9213102-A.
 PD 06-AUG-1992.
 PE 15-JAN-1992; U00340.
 PR 15-JAN-1991; US-642342.
 PA (GENM-) GENMARK.
 PI Georges M, Massey JM;
 DR WPI: 92-284684/34.
 PT Polymorphic bovine DNA markers - used in genetic identification,
 PS gene mapping, and selective breeding
 Table 7: Page 282: 517pp: English.
 CC The sequence is that of a bovine microsatellite sequence obtd. by
 CC screening a library of bovine MboI DNA fragments of between
 CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
 CC One out of 50 clones cross-hybridised. Assuming independent
 CC distribution of microsatellites and MboI sites, the frequency of
 CC (T6)n > 9 microsatellites in the bovine genome is estimated at >100,
 CC 000. The sequence information for ca. 230 such bovine microsatellites
 CC is summarised in the specification and indexed herein (see below).
 CC The sequences upstream and downstream of the microsatellite sequence
 CC were used to generate the required PCR primers for in vitro
 CC amplification of the corresp. microsatellite (using the program
 CC OPTIPRIM). The microsatellites may be used to identify individuals,
 CC for parentage testing, and in the genetic mapping of economic trait
 CC loci, or genes involved in the determination of economically important
 CC traits esp. in cattle, to allow selective breeding.
 CC See also Q33501-34437.
 SQ Sequence 36 BP; 0 A; 18 G; 18 T;

Query Match 0.1%; Score 36; DB 1; Length 36;
 Best Local Similarity 100.0%; Pred. No. 0.0031;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CACACACACACACACACACACACACACA 36
 Db 36 CACACACACACACACACACACACACACA 1

RESULT 7
 ID Q33953 standard; DNA; 36 BP.
 AC Q33953;
 DT 02-FEB-1993 (first entry)
 DE Microsatellite sequence from clone TGLA35.
 KW PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;
 KM genetic mapping; traits; amplification; ss.
 OS Bos taurus.
 PN WO9213102-A.
 PD 06-AUG-1992.
 PE 15-JAN-1992; U00340.
 PR 15-JAN-1991; US-642342.
 PA (GENM-) GENMARK.
 PI Georges M, Massey JM;
 DR WPI: 92-284684/34.
 PT Polymorphic bovine DNA markers - used in genetic identification,
 PS gene mapping, and selective breeding
 Table 7: Page 311: 517pp: English.
 CC The sequence is that of a bovine microsatellite sequence obtd. by
 CC screening a library of bovine MboI DNA fragments of between
 CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
 CC One out of 50 clones cross-hybridised. Assuming independent
 CC distribution of microsatellites and MboI sites, the frequency of
 CC (T6)n > 9 microsatellites in the bovine genome is estimated at >100,
 CC 000. The sequence information for ca. 230 such bovine microsatellites
 CC is summarised in the specification and indexed herein (see below).
 CC The sequences upstream and downstream of the microsatellite sequence
 CC were used to generate the required PCR primers for in vitro
 CC amplification of the corresp. microsatellite (using the program
 CC OPTIPRIM). The microsatellites may be used to identify individuals,
 CC for parentage testing, and in the genetic mapping of economic trait
 CC loci, or genes involved in the determination of economically important
 CC traits esp. in cattle, to allow selective breeding.
 CC See also Q33501-34437.
 SQ Sequence 36 BP; 0 A; 18 G; 18 T;

Query Match 0.1%; Score 36; DB 1; Length 36;
 Best Local Similarity 100.0%; Pred. No. 0.0031;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CACACACACACACACACACACACACACA 36
 Db 36 CACACACACACACACACACACACACACA 1

RESULT 8
 ID Q34041 standard; DNA; 37 BP.
 AC Q34041;
 DT 02-FEB-1993 (first entry)
 DE Microsatellite sequence from clone TGLA433.
 KW PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;
 KM genetic mapping; traits; amplification; ss.
 OS Bos taurus.
 PN WO9213102-A.
 PD 06-AUG-1992.
 PE 15-JAN-1992; U00340.
 PR 15-JAN-1991; US-642342.
 PA (GENM-) GENMARK.
 PI Georges M, Massey JM;
 DR WPI: 92-284684/34.
 PT Polymorphic bovine DNA markers - used in genetic identification,

PS Table 7; page 189; 517bp. Bovine microsatellite sequence obt'd. by
CC The sequence is that of a bovine microsatellite sequence obtained by
CC screening a library of bovine MboI DNA fragments of between
CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
CC One out of 50 clones cross-hybridised, assuming independent
CC distribution of microsatellites and MboI sites, the frequency of
CC (Tn)₉ microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for *in vitro*
CC amplification of the corresp. microsatellite (using the program
CC OPTIRN1). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved in the determination of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SQ Sequence 39 BP, 0 A, 0 C, 19 G, 20 T,

Query Match	0.18;	Score 36;	DB 1;	Length 39;
Best Local Similarity	100.08;	Pred. No. 0.003;		
Matches 36;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0.

QY	1	CACACACACACACACACACACACACACACA	36
Db	38	CACACACACACACACACACACACACACACA	3

RESULT 15
 ID 033807/C
 AC 033807:
 DT 02-FEB-1993 (first entry)
 DE Microsatellite sequence from clone TGLA213.
 KW PCR; selection; primers; OPR191M; breeding; cattle; parentage;
 KW genetic mapping; traits; amplification; ss.
 OS Bos taurus.
 PN M09213102-A.
 PD 06-AUG-1992.
 PF 15-JAN-1992;
 PI 15-JAN-1991; US-642342.
 PR (GENM-7) GENMARK.
 PI Georges W, Massey JM;
 DR WJL; 92-284684/34.
 PT Polymorphic bovine DNA markers - used in genetic identification,
 PT gene mapping, and selective breeding.
 PS Table 7, Page 252; 51pp; English.
 CC The sequence is that of a bovine microsatellite sequence obtd. by
 CC screening a library of bovine MboI DNA fragments of between
 CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
 CC One out of 50 clones cross-hybridised. Assuming independent
 CC distribution of microsatellites and MboI sites, the frequency of
 CC (76)_n > microsatellites in the bovine genome is estimated at >100,
 CC 000. The sequence information for ca. 230 such bovine microsatellites
 CC is summarised in the specification and indexed herein (see below).
 CC The sequences upstream and downstream of the microsatellite sequence
 CC were used to generate the required PCR primers for in vitro
 CC amplification of the corresp. microsatellite (using the program
 CC OPTIRIV). The microsatellites may be used to identify individuals,
 CC for parentage testing, and in the genetic mapping of economic trait
 CC loci or genes involved in the determinism of economically important
 CC traits esp. in cattle, to allow selective breeding.
 CC See also Q33501-34437.
 SO Sequence 39 Bp; 0 A; 0 C; 20 G; 19 T;

```
Query Match          0 1%; Score 36; DB 1; Length 39;  
Best Local Similarity 100.0%; Pred.No. 0.003;  
Matches    36; Conservative   0; Mismatches   0; Indels   0  
  
      1 CACACACACACACACACACACACACACACA 36  
      |||
```

Db 39 CACACACACACACACACACACACACACACACACACACA
Search completed: March 29, 2000, 13:08:06
Job time: 33484 sec

Mon Apr 3 08:24:16 2000

us-08-852-495c-1_copy_1_30000.rng

Page 7



9
v
4

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 10:09:37 ; Search time 8959.53 Seconds

(without alignments)
-6778.337 Million cell updates/sec

Title: US-08-852-495c-1_COPY_25000_45000

Perfect score: 20001
Sequence: 1 AGTGCAGTACTGCGACCTC.....CAAGTTAGAGACATTTT 20001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 821193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : GenDbml:*

1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pr1:*
11: gb_pr2:*
12: gb_pr3:*
13: gb_ro:*
14: gb_sts:*
15: gb_sy:*
16: gb_un:*
17: gb_v1:*
18: em_fun:*
19: em_hum1:*
20: em_hum2:*
21: em_hum3:*
22: em_om:*
23: em_or:*
24: em_ov:*
25: em_pat:*
26: em_ph:*
27: em_pl:*
28: em_ro:*
29: em_sy:*
30: em_un:*
31: em_v1:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*

44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	89	0.4	96	13	G31304	G31304 sy89991-19
2	48	0.2	56	5	I31237	I31237 Sequence 14
3	46	0.2	85	10	HUMHIS1PR	M26162 Homo sapien
4	43	0.2	54	5	AR002288	AR002288 Sequence
5	43	0.2	54	5	AR053139	AR053139 Sequence
6	42	0.2	69	13	HUMUT578A	L39139 Human STS U
7	40	0.2	40	5	A68621	A68621 Sequence 1
8	40	0.2	101	10	S79560	S79560 HRX (part1
9	40	0.2	101	10	S79561	S79561 dhrx (part1
10	39	0.2	39	5	I31174	I31174 Sequence 86
11	39	0.2	39	5	I82155	I82155 Sequence 1
12	39	0.2	75	13	AF042304	AF042304 Homo sapl
13	39	0.2	102	39	SHPDINCRD	L01534 Sheep dlnuc
14	38	0.2	60	5	AR051487	AR051487 Sequence
15	38	0.2	66	5	AR002280	AR002280 Sequence
16	38	0.2	66	5	AR053141	AR053141 Sequence
17	38	0.2	76	5	AR051499	AR051499 Sequence
18	37	0.2	59	9	HSDBS30A3	X54497 H. sapiens 3
19	37	0.2	73	13	AF042303	AF042303 Homo sapl
20	36	0.2	40	5	I31240	I31240 Sequence 15
21	36	0.2	66	40	AF087511	AF087511 Homo sapl
22	35	0.2	35	5	A25212	A25212 Inter-Alu s
23	35	0.2	35	24	E09140	E09140 Synthetic D
24	35	0.2	45	5	I31245	I31245 Sequence 15
25	35	0.2	80	9	HUMBRKPAD	M36134 Human alpha
26	35	0.2	80	9	HUMBRKPAE	M36135 Human alpha
27	35	0.2	85	5	AR051522	AR051522 Sequence
28	34	0.2	40	5	A68621	A68621 Sequence 1
29	34	0.2	70	13	HUMUT5148A	L31067 Human STS U
30	34	0.2	87	12	MWIVW27	X89144 M.musculus
31	33	0.2	35	5	A25212	A25212 Inter-Alu s
32	33	0.2	35	24	E09140	E09140 Synthetic D
33	33	0.2	55	9	HSDBS30A1	X54495 H. sapiens 3
34	33	0.2	70	12	MUSMABA02	K02420 Mouse MHC c
35	33	0.2	80	5	HUMBRKPAE	M36133 Human alpha
36	33	0.2	83	13	A08915	A08915 H. sapiens (
37	32	0.2	90	9	AU025339	AU025339 Rattus no
38	32	0.2	90	13	HUMCDLREL	K03555 Human low d
39	32	0.2	90	13	G37815	G37815 ACT1 Plasmid
40	32	0.2	100	11	HS067848	U67848 Human beta-
41	32	0.2	102	5	A08911	A08911 H. sapiens (
42	32	0.2	104	13	AU046973	AU046973 Rattus no
43	31	0.2	60	11	HSCBPS108	AF084962 Homo sapl
44	31	0.2	62	10	S52155S31	S52228 CD11b-1enko
45	31	0.2	71	13	HUMUT741A	L29638 Human STS U

ALIGNMENTS

RESULT 1

G31304 96 bp DNA STS 29-SEP-1998
LOCUS sy89991-19 Human (A.Guilke) Homo sapiens STS genomic, sequence
DEFINITION tagged site.
ACCESSION G31304

[illegible]

```

TITLE
JOURNAL
COMMENT
Melis, R., Lawrence, E., Moore, M., Holik, P.R., Carlson, M., Zhao, X.,
Robertson, M., Bradley, P., Elsner, T., Tingey, A., Lalouel, J.-M. and
White, R.
Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
Unpublished (1994)
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: sts@corona.med.utah.edu
Primer A: AGTTCGAGACCGACCTGGC
Primer B: AGGTGGCAGAAATCGCATC
End to label: Primer A
PCR profile:
Initial Denaturation: 94C 300sec
Cycles Denaturation Annealing Extension 5 94
C 10 sec. 62 C 10 sec. 72 C 20 sec. 30
58 C 10 sec. 72 C 20 sec. Mg++: 1.50 mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 1.
1. location/Qualifiers
1. 69
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19"
14. 32
/note="STS UT578 5' end"
/evidence=experimental
26 a 16 c 14 g 13 t
BASE COUNT
ORIGIN

```

Query Match	0.2%	Score 42	DB 13	Length 69
Best Local Similarity	100.0%	Pred. No. 3	7e-10	
Matches 42	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 10109	TTGAGTCAAGAGATTGAGACCAAGCCCTGGCCCAACATGTGTAA	10150		
DB 3	TTGAGTCAAGAGATTGAGACCAAGCCCTGGCCCAACATGTGTAA	44		
RESULT 7				
A68621/c				
LOCUS	A68621	40 bp	DNA	PAT
DEFINITION	Sequence 1 from Patent WO9801573.			06-MAY-1999
ACCESSION	A68621			
VERSION	A68621.1	GI:4759648		
KEYWORDS				
SOURCE	unidentified.			
ORGANISM	unidentified			
REFERENCE	unclassified.			
AUTHORS	1 (bases 1 to 40)			
TITLE	Resnick, M.A., Laktionov, V.L., Kouprina, N.Y. and Perkins, E.L.			
JOURNAL	TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING			
	Patent: WO 9801573-A 15-JAN-1998;			
	US HEALTH (US)			
FEATURES				
source	1..40	location/Qualifiers		
	/organism="unidentified"			
	/db_xref="taxon:32644"			
BASE COUNT	7 a 12 c 13 g	8 t		
ORIGIN				
Query Match	0.2%	Score 40	DB 5	Length 40
Best Local Similarity	100.0%	Pred. No. 4	1e-09	
Matches 40	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 8640	TCGGCCTCCCAAGATGCTGGATTCAGAGCGTGAGCCACC	8679		
DB 40	TCGGCCTCCCAAGATGCTGGATTCAGAGCGTGAGCCACC	1		

```
RESULT 8
LOCUS 579560 101 bp DNA PRI 27-JAN-1996
DEFINITION HRX [intron 6] [human, genomic, 101 nt].
ACCESSION 579560
VERSION 579560.1 GI:1168041
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Bernard,O.A., Romana,S.P., Schichman,S.A., Mauchauffe,M.,
Jonveaux,P. and Berger,R.
TITLE Partial duplication of HRX in acute leukemia with trisomy 11
JOURNAL Leukemia 9 (9), 1487-1490 (1995)
MEDLINE 95387660
REMARK Genbank staff at the National Library of Medicine created this
entry [NCBI gisbseq 170933] from the original journal article.
This sequence comes from Fig. 3.
Map location: 11q23.
FEATURES
source location/Qualifiers
1..101
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..101
/Partial
/gene="HRX"
BASE COUNT 27 a 21 c 28 g 25 t
ORIGIN

Query Match 0.2%; Score 40; DB 10; Length 101;
Best Local Similarity 100.0%; Pred. No. 4e-09;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 12877 GAGACGAGCCTGGCCACATGTGTGAACCTGTCTCTACT 12916
Db 11 GAGACGAGCCTGGCCACATGTGTGAACCTGTCTCTACT 50
|||||
RESULT 9
LOCUS 579561 101 bp DNA PRI 27-JAN-1996
DEFINITION HRX [partial genomic duplication startpoint] [human, acute myeloid
leukemia with trisomy 11 syndrome patient J, genomic Mutant, 101
nt].
ACCESSION 579561
VERSION 579561.1 GI:1168042
KEYWORDS
SOURCE human acute myeloid leukemia with trisomy 11 syndrome patient J.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Bernard,O.A., Romana,S.P., Schichman,S.A., Mauchauffe,M.,
Jonveaux,P. and Berger,R.
TITLE Partial duplication of HRX in acute leukemia with trisomy 11
JOURNAL Leukemia 9 (9), 1487-1490 (1995)
MEDLINE 95387660
REMARK Genbank staff at the National Library of Medicine created this
entry [NCBI gisbseq 170936] from the original journal article.
This sequence comes from Fig. 3.
Map location: 11q23.
Authors indicate breakpoint region at 48-62.
FEATURES
source location/Qualifiers
1..101
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..101
/Partial
/gene="HRX"
BASE COUNT 12 a 8 c 15 g 4 t
ORIGIN

Query Match 0.2%; Score 39; DB 5; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 16299 CTTGGTCTCTCACTGACATGATATCTGTCTTCC 16337
Db 11 CTTGGTCTCTCACTGACATGATATCTGTCTTCC 16337
|||||
```


[illegible]

```

/db_xref="taxon:9940"
/cell_type="leukocyte"
repeat_region
31..69
/standard_name="dinucleotide repeat"
BASE COUNT      40 a      42 c      9 g      11 t
ORIGIN

Query Match          0.2%; Score 39; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   7517 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7555
      |||||||
Db    69 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 31

RESULT 14
AR051487/c        AR051487     60 bp     DNA       PAT       29-SEP-1999
DEFINITION        Sequence 57 from patent US 5830670.
ACCESSION         AR051487
VERSION           AR051487.1 GI:5974851
KEYWORDS
SOURCE            Unknown.
ORGANISM          Unclassified.
REFERENCE         1 (bases 1 to 60)
AUTHORS           de la Monte,S. and Wands,J.R.
TITLE             Neural thread protein gene expression and detection of Alzheimer's
                  disease
JOURNAL           Patent: US 5830670-A 57 03-NOV-1998;
FEATURES          Location/Qualifiers
                    ..60
                    /organism="unknown"
BASE COUNT        12 a      14 c      15 g      19 t
ORIGIN

Query Match          0.2%; Score 38; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   12897 GGTGAACCCGTCTCTACTATAAAAAATACAAAATTAGC 12934
      |||||||
Db    48 GGTGAACCCGTCTCTACTATAAAAAATACAAAATTAGC 11

RESULT 15
AR002290          AR002290     66 bp     DNA       PAT       04-DEC-1998
DEFINITION        Sequence 29 from patent US 5741645.
ACCESSION         AR002290
VERSION           AR002290.1 GI:3963844
KEYWORDS
SOURCE            Unknown.
ORGANISM          Unclassified.
REFERENCE         1 (bases 1 to 66)
AUTHORS           Orr,H.T., Rannum,L.P.W., Chung,M. and Zoghbi,H.Y.
TITLE             Gene sequence for spinocerebellar ataxia type I and method for
                  diagnosis
JOURNAL           Patent: US 5741645-A 29 21-APR-1998;
FEATURES          Location/Qualifiers
                    ..66
                    /organism="unknown"
BASE COUNT        12 a      0 c      31 g      23 t
ORIGIN

Query Match          0.2%; Score 38; DB 5; Length 66;
Best Local Similarity 100.0%; Pred. No. 4.3e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Mon Apr 3 08:24:22 2000

us-08-852-495c-1_copy_25000_45000.rge

Page 6

```
Oy    7513 GTGTGCTGTGTGTGTGTGTGTATGTGTGTGT 7550  
      |||||  
Db    29 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 66
```

Search completed: March 29, 2000, 17:18:26
Job time: 50055 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 17:18:26 ; Search time 13410.2 Seconds
(without alignments)
-6792.946 Million cell updates/sec

Title: US-08-852-495C-1_COPY_40000_70000
Perfect score: 30001
Sequence: 1 GGTGACACGCTGACGTA.....AACACAGTGCAGATTGA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 821193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_da1:*
2: gb_da2:*
3: gb_da3:*
4: gb_da4:*
5: gb_da5:*
6: gb_da6:*
7: gb_da7:*
8: gb_da8:*
9: gb_da9:*
10: gb_da10:*
11: gb_da11:*
12: gb_da12:*
13: gb_da13:*
14: gb_da14:*
15: gb_da15:*
16: gb_da16:*
17: gb_da17:*
18: gb_da18:*
19: gb_da19:*
20: gb_da20:*
21: gb_da21:*
22: gb_da22:*
23: gb_da23:*
24: gb_da24:*
25: gb_da25:*
26: gb_da26:*
27: gb_da27:*
28: gb_da28:*
29: gb_da29:*
30: gb_da30:*
31: gb_da31:*
32: gb_da32:*
33: gb_da33:*
34: gb_da34:*
35: gb_da35:*
36: gb_da36:*
37: gb_da37:*
38: gb_da38:*
39: gb_da39:*
40: gb_da40:*
41: gb_da41:*
42: gb_da42:*
43: gb_da43:*

44: gb_da44:*
45: gb_da45:*
46: gb_da46:*
47: gb_da47:*
48: gb_da48:*
49: gb_da49:*
50: gb_da50:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	89	0.3	96	13 G31304	G31304 sy899g1-19
2	46	0.2	92	8 AF062763	AF062763 Glycine m
3	45	0.1	99	13 G37948	G37948 C4W78 plasm
4	44	0.1	99	13 G37948	G37948 C4W78 plasm
5	43	0.1	90	9 HUMDLRPM	M15365 Human low d
6	42	0.1	67	9 HUMALDANC	L36843 Homo sapien
7	41	0.1	90	9 HUMDLRPM	M15365 Human low d
8	41	0.1	92	8 AF062763	AF062763 Glycine m
9	41	0.1	104	9 HUMALCE272	M67899 Human carc1
10	39	0.1	39	5 I82155	I82155 Sequence 1
11	39	0.1	60	5 I88893	I88893 Sequence 11
12	39	0.1	80	9 HUMBRKFAA	M36131 Human alpha
13	38	0.1	40	5 A68622	A68622 Sequence 2
14	38	0.1	58	5 AR063068	AR063068 Sequence
15	38	0.1	62	10 S52152531	S52228 CD11b-Leuko
16	38	0.1	70	5 A49117	A49117 Sequence 12
17	38	0.1	75	5 A49118	A49118 Sequence 13
18	38	0.1	80	5 A18777	A18777 5' terminus
19	38	0.1	80	5 AR001312	AR001312 Sequence
20	38	0.1	90	9 HUMDLRFL	K03555 Human low d
21	37	0.1	37	5 I29931	I29931 Sequence 44
22	37	0.1	40	5 A48799	A48799 Sequence 6
23	37	0.1	50	5 I36502	I36502 Sequence 1
24	37	0.1	55	5 I29928	I29928 Sequence 41
25	37	0.1	69	5 I89934	I89934 Sequence 7
26	37	0.1	69	5 I89944	I89944 Sequence 22
27	37	0.1	71	10 S76508	S76508 DIS8 (B) fh
28	37	0.1	72	10 S76510	S76510 DIS8 (B) fh
29	37	0.1	73	5 A08915	A08915 H. sapiens (
30	37	0.1	77	5 A08908	A08908 H. sapiens (
31	37	0.1	80	5 A08907	A08907 H. sapiens (
32	37	0.1	80	9 HUMBRKFAA	M36132 Human alpha
33	37	0.1	80	9 HUMBRKFAA	M36133 Human alpha
34	37	0.1	80	9 HUMBRKFAA	M36133 Human alpha
35	37	0.1	84	5 AR038854	AR038854 Sequence
36	37	0.1	84	5 AR050387	AR050387 Sequence
37	37	0.1	90	5 E12579	E12579 Probe. 6/19
38	37	0.1	90	5 E12580	E12580 Probe. 6/19
39	37	0.1	90	5 E13639	E13639 Probe. 6/19
40	37	0.1	90	5 E17120	E17120 DNA probe.
41	37	0.1	91	34 DDIACTIN8A	M25215 Dictyostell1
42	37	0.1	96	5 A08909	A08909 H. sapiens (
43	37	0.1	98	5 I49625	I49625 Sequence 42
44	37	0.1	102	5 A08911	A08911 H. sapiens (
45	37	0.1	104	5 A08910	A08910 H. sapiens (

ALIGNMENTS

RESULT 1
LOCUS G31304 96 bp DNA
DEFINITION sy899g1-19 Human (A.Gnlrke) Homo sapiens STS genomic, sequence
ACCESSION G31304
tagged site.

Primer: 5 pM each
dNTPs: 0.3 uL of 10 mM stock solution
Tag Polymerase: 0.5 unit
Total Volume: 15 uL

Buffer: 10 mM Tris HCl, pH 8.3
1.5 mM MgCl2
50 mM KCl.

FEATURES
source
Location/Qualifiers
1. .99
/organism="Plasmodium falciparum"
/strain="HB3xDd2"
/db_xref="taxon:5833"
/map="4"
/clone_lib="Plasmodium falciparum haploid"
/dev_stage="erythrocytic haploid stage"

STS
primer_bind 1. .99
primer_bind 1. .99
complement(79..99)

BASE COUNT 47 a 6 c 10 g 36 t
ORIGIN

Query Match 0.1%; Score 45; DB 13; Length 99;
Best Local Similarity 100.0%; Pred. No. 4.9e-12;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20223 TATGA 20267
Db 63 TATGA 19

RESULT 4
G37948 99 bp DNA STS 26-FEB-1999
LOCUS CAM78 plasmodium falciparum haploid Plasmodium falciparum STS
DEFINITION genomic, sequence tagged site.
ACCESSION G37948
VERSION G37948.2 GI:4337425
KEYWORDS STS.
SOURCE malaria parasite P. falciparum.
ORGANISM Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 99)
AUTHORS Su, X.Z., Ferdig, M.T. and Wellens, T.E.
TITLE Integrated genetic and physical maps of Plasmodium falciparum
JOURNAL Unpublished (1998)
COMMENT On Mar 6, 1999 this sequence version replaced gi:3056692.

Contact: Thomas E. Wellens

LPD
NIAID
Bldg. 4, Room 126, NIH Campus, Bethesda, MD 20892-0425
Tel: (301) 496-4021
Fax: (301) 402-0079
Email: tewhell@nih.gov
Primer A: GGAATATACACAGCAATC
Primer B: AATTACTCTGCAATTTC
STS size: 99
PCR Profile:
Initial heat: 2min at 94°C
Cycles: 30 cycles, each having steps 20sec at 94°C;
10sec at 45°C; 10sec at 40°C; 30sec at 60°C
Machine: Perkin Elmer GeneAmp 9600
Protocol:
Template: 1-10 ng
Primer: 5 pM each
dNTPs: 0.3 uL of 10 mM stock solution
Tag Polymerase: 0.5 unit
Total Volume: 15 uL

Buffer:

10 mM Tris HCl, pH 8.3
1.5 mM MgCl2
50 mM KCl.

FEATURES
source
Location/Qualifiers
1. .99
/organism="Plasmodium falciparum"
/strain="HB3xDd2"
/db_xref="taxon:5833"
/map="4"
/clone_lib="Plasmodium falciparum haploid"
/dev_stage="erythrocytic haploid stage"

STS
primer_bind 1. .99
primer_bind 1. .99
complement(79..99)

BASE COUNT 47 a 6 c 10 g 36 t
ORIGIN

Query Match 0.1%; Score 44; DB 13; Length 99;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20223 TATG 20266
Db 22 TATG 65

RESULT 5
HUMDLRM 90 bp DNA PRI 11-JAN-1995
LOCUS Human low density lipoprotein receptor mutant gene recombination
DEFINITION site.
ACCESSION M15365
VERSION M15365.1 GI:187107
KEYWORDS Alu repeat; LDL receptor; cell surface protein; crossover;
SOURCE recombination.
ORGANISM Human (FH 295) fibroblast DNA, clone p295.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 90)
AUTHORS Lehman, M.A., Goldstein, J.L., Russell, D.W. and Brown, M.S.
TITLE Duplication of seven exons in LDL receptor gene caused by Alu-Alu
JOURNAL recombination in a subject with familial hypercholesterolemia
MEDLINE Cell 48 (5), 827-835 (1987)
COMMENT Clean copy of sequence kindly provided by M. Lehman (22-APR-1987).
Individual FH 295 carries two mutant LDL receptor alleles. The
allele below was inherited from the father and includes a
duplication of exons 2 through 8 that was presumably created by
unequal chromosomal crossover involving Alu repeats in introns 1
and 8. The exact site of recombination cannot be determined
because the sequences of the Alu repeats in introns 1 and 8 of
normal alleles are identical over a span of 26 nucleotides at the
recombination site.

FEATURES
source
Location/Qualifiers
1. .90
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19p13.3"
1. .43
/gene="LDLR"
1. .43
/gene="LDLR"
/note="LDLR intron 8; G00-119-362"
43. .68
/organism="Homo sapiens"
68. .>90
/note="LDLR duplicated intron 1 (no splice consensus at
68); putative; does not fit consensus"

BASE COUNT 18 a 33 c 19 g 20 t
ORIGIN 1 bp upstream of NheI site; chromosome 19p13.2-p13.1.

Query Match 0.1%; Score 43; DB 9; Length 90;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9929 CCTCAGCTCCCAAGTGGTGGATTACAGCATGACCCACCA 9971
|||||
DB 38 CCTCAGCTCCCAAGTGGTGGATTACAGCATGACCCACCA 80

RESULT 6
HUMALUNCD 67 bp DNA PRI 08-OCT-1994
LOCUS Homo sapiens 4000 year old remains from Nekht-ankh Alu repeat
DEFINITION fragment 12:5.
ACCESSION I36843.1 GI:556196
VERSION I36843
KEYWORDS Alu repeat.
SOURCE Homo sapiens (individual isolate 4000 year old remains from Nekht-ankh) liver DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Pabho, S
TITLE Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (6), 1939-1943 (1989)
MEDLINE 89184542
FEATURES
Source location/Qualifiers
1..67
/organism="Homo sapiens"
/isolate="4000 year old remains from Nekht-ankh"
/db_xref="taxon:9606"
/tissue_type="liver"
/tissue_type="liver"
repeat_region /rpt_family="Alu"
BASE COUNT 15 a 18 c 23 g 11 t
ORIGIN

Query Match 0.1%; Score 42; DB 9; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20914 GTGGCTCAGCTGTAATCCAGCACTTGGAGGCTGAGGC 20955
|||||
DB 9 GTGGCTCAGCTGTAATCCAGCACTTGGAGGCTGAGGC 50

RESULT 7
HUMDLRM/c 90 bp DNA PRI 11-JAN-1995
LOCUS Human low density lipoprotein receptor mutant gene recombination site.
DEFINITION
ACCESSION M15365.1 GI:187107
VERSION M15365.1
KEYWORDS Alu repeat; LDL receptor; cell surface protein; crossover; recombination.
SOURCE Human (FH 295) fibroblast DNA, clone p295.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 90)
AUTHORS Lehman, M.A., Goldstein, J.L., Russell, D.W. and Brown, M.S.
TITLE Duplication of seven exons in LDL receptor gene caused by Alu-Alu recombination in a subject with familial hypercholesterolemia
JOURNAL Cell 48 (5), 827-835 (1987)
MEDLINE 87131094
COMMENT Clean copy of sequence kindly provided by M. Lehman (22-APR-1987). Individual FH 295 carries two mutant LDL receptor alleles. The allele below was inherited from the father and includes a duplication of exons 2 through 8 that was presumably created by

unequal chromosomal crossover involving Alu repeats in introns 1 and 8. The exact site of recombination cannot be determined because the sequences of the Alu repeats in introns 1 and 8 of normal alleles are identical over a span of 26 nucleotides at the recombination site

FEATURES
Source location/Qualifiers
1..90
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19p13.3"
1..43
/gene="LDLR"
/note="LDLR"
/gene="LDLR"
/note="LDLR intron 8; G00-119-362"
43..68
/organism="Homo sapiens"
68..>90
/note="LDLR duplicated intron 1 (no splice consensus at 68); putative; does not fit consensus"
18 a 19 g 20 t
BASE COUNT 18 a 33 c 19 g 20 t
ORIGIN 1 bp upstream of NheI site; chromosome 19p13.2-p13.1.

Query Match 0.1%; Score 41; DB 9; Length 90;
Best Local Similarity 100.0%; Pred. No. 6.3e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23593 GTGGCTCAGCTGTAATCCAGCACTTGGAGGCTGAGG 23633
|||||
DB 78 GTGGCTCAGCTGTAATCCAGCACTTGGAGGCTGAGG 38

RESULT 8
AF062763/c 92 bp DNA PLN 02-AUG-1998
LOCUS Glycine max microsatellite AC21 repeat region.
DEFINITION
ACCESSION AF062763
VERSION AF062763.1 GI:3372757
KEYWORDS soybean.
SOURCE Glycine max
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
REFERENCE 1 (bases 1 to 92)
AUTHORS Peakall, R., Gilmore, S., Keys, W., Morgante, M. and Rafalski, A.
TITLE Cross-species amplification of soybean (Glycine max) simple-sequence-repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants
JOURNAL Mol. Biol. Evol. (1998) in press
2 (bases 1 to 92)
REFERENCE Peakall, R., Gilmore, S., Keys, W., Morgante, M. and Rafalski, A.
TITLE Direct Submision
JOURNAL Submitted (04-MAY-1998) Division of Botany and Zoology, Australian National University, Canberra, ACT 0200, Australia
FEATURES
Source location/Qualifiers
1..92
/organism="Glycine max"
/db_xref="taxon:3847"
repeat_region /note="microsatellite AC21"
/rpt_type="tandem"
/rpt_unit="ta 5 c 4 g 35 t
BASE COUNT 48 a 5 c 4 g 35 t
ORIGIN

Query Match 0.1%; Score 41; DB 8; Length 92;
Best Local Similarity 100.0%; Pred. No. 6.3e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Mon Apr 3 08:24:25 2000

us-08-852-495c-1_copy_40000_70000.rge

Page 6

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
BASE COUNT
ORIGIN

unidentified.
unidentified.
unclassified.
1 (bases 1 to 40)
Resnick, M.A., Laktionov, V.L., Koupina, N.Y. and Perkins, E.L.
TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
Patent: WO 9801573-A 15-JAN-1998;
US HEALTH (US)
Location/Qualifiers
1..40
/organism="unidentified"
/db_xref="taxon:32644"
9 a 8 c 19 g 4 t

Query Match 0.1%; Score 38; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8613 GAGGCTGAGCAGAGAACGCTGACCGGAGCGG 8650
|||||
Db 2 GAGGCTGAGCAGAGAACGCTGACCGGAGCGG 39

RESULT 14
LOCUS AR063068 58 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 42 from patent US 5844075.
ACCESSION AR063068
VERSION AR063068.1 GI:5990759
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 58)
Kawakami, Y. and Rosenberg, S.A.
TITLE Melanoma antigens and their use in diagnostic and therapeutic methods
JOURNAL Patent: US 5844075-A 42 01-DEC-1998;
FEATURES Location/Qualifiers
1..58
SOURCE /organism="unknown"
BASE COUNT 3 a 6 c 9 g 40 t
ORIGIN

Query Match 0.1%; Score 38; DB 5; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13600 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 13637
|||||
Db 18 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 55

RESULT 15
LOCUS S52152531/c 62 bp DNA PRI 08-MAY-1993
DEFINITION CD11b-leukocyte integrin alpha chain [human, Genomic, 62 nt,
segment 31 of 31].
ACCESSION S52228
VERSION S52228.1 GI:263048
KEYWORDS
SEGMENT 31 of 31
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 62)
Fleming, J.C., Pahl, H.L., Gonzalez, D.A., Smith, T.F. and Tenen, D.G.
TITLE Structural analysis of the CD11b gene and phylogenetic analysis of

the alpha-integrin gene family demonstrate remarkable conservation
of genomic organization and suggest early diversification during
evolution
JOURNAL U. Immunol. 150 (2), 480-490 (1993)
MEDLINE 93123748
REMARK Genbank staff at the National Library of Medicine created this
entry [NCBI g1dbsg 121959] from the original journal article.
This sequence comes from Fig. 3.
Map location: 16.
Location/Qualifiers
1..62
/organism="Homo sapiens"
/db_xref="taxon:9606"
51 a 2 c 1 g 8 t

Query Match 0.1%; Score 38; DB 10; Length 62;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13601 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 13638
|||||
Db 58 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 21

Search completed: March 30, 2000, 00:22:34
Job time: 75303 sec

Mon Apr 3 08:24:25 2000

us-08-852-495c-1_copy_40000_70000.rge



2

7

•

CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/341,562
 FILING DATE: 21-APR-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Sara, Charles S.
 REGISTRATION NUMBER: 30,492
 REFERENCE/DOCKET NUMBER: 09865.601
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (608) 831-2100
 TELEFAX: (608) 831-2106
 TELEX:
 INFORMATION FOR SEQ ID NO: 198:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 38 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 IMMEDIATE SOURCE:
 CLONE: mfd50tS
 US-08-222-177A-198

	Query Match	0.1%	Score 36;	DB 1;	Length 38;
	Best Local Similarity	100.0%	Pred. No.	0.00084;	
Matches	36;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
OY	1 CACACACACACACACACACACACACACA	36			
Db	1 CACACACACACACACACACACACACACA	36			

```

1      RESULT      5
2      US-08-222-177A-397
3      ; Sequence 397, Application US/08222177A
4      ; Patent No. 3582979
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Weber, James L.
7      ; TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
8      ; TITLE OF INVENTION: (dC-dA)n..(dG-dT)n SEQUENCES AND METHODS OF USING SAME
9      ; NUMBER OF SEQUENCES: 460
10     ; CORRESPONDENCE ADDRESSES:
11     ; ADDRESSEE: Dewitt Ross & Stevens, S.C.
12     ; STREET: 8000 Excelsior Drive, Suite 401
13     ; CITY: Madison
14     ; STATE: Wisconsin
15     ; COUNTRY: USA
16     ; ZIP: 53717-1914
17     ; COMPUTER READABLE FORM:
18     ; MEDIUM TYPE: Floppy disk
19     ; COMPUTER: IBM PC compatible
20     ; OPERATING SYSTEM: PC-DOS/MS-DOS
21     ; SOFTWARE: PatentIn Release #1.0, Version #1.25
22     ; CURRENT APPLICATION DATA:
23     ; APPLICATION NUMBER: US/08/222,177A
24     ; FILING DATE:
25     ; CLASSIFICATION: 435
26     ; PRIOR APPLICATION DATA:
27     ; APPLICATION NUMBER: US 07/341,562
28     ; FILING DATE: 21-APR-1989
29     ; ATTORNEY/AGENT INFORMATION:
30     ; NAME: Sara, Charles S.
31     ; REGISTRATION NUMBER: 30,492
32     ; REFERENCE/DOCKET NUMBER: 09865, 601
33     ; TELECOMMUNICATION INFORMATION:
34     ; TELEPHONE: (608) 831-2100
35     ; TELEFAX: (608) 831-2106
36     ; TELEX:
37     ; INFORMATION FOR SEQ ID NO.: 397:
38     ; SEQUENCE CHARACTERISTICS:
39     ; LENGTH: 38 base pairs
40     ; TYPE: nucleic acid
41     ;

```

```

; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-222-177A-397

```

```

Query Match      0.1%; Score 36; DB 1; Length 38;
Best Local Similarity 100.0%; Pred. No. 0 00084;
Matches 36; Conservative 0; Mismatches 0; Gaps 0;

```

Qy 1 CACACACACACACACACACACACACACACA 36
 |||||
Db 2 CACACACACACACACACACACACACACACA 37

RESULT 6
US-08-222-177A-137
Sequence 137, Application US/08222177A
Date of US Filing 08/08/2008

```

: APPLICANT: Weber, James L.
: TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
: TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
: TITLE OF INVENTION:
:

```

ADDRESS: DEWITT ROSS & STEVENS, S.C.
STREET: 8000 EXCELSIOR DRIVE, SUITE 401
CITY: MADISON

COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy

```

:
: OF ENVIRONMENTAL STATEMENTS:
: SOFTWARE: Patent Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/222,177A
:

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989

NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION.

TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 137:
SEQUENCE CHARACTERISTICS:

```

;
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;

```

IMMEDIATE SOURCE:
CLONE: mfd29rs
US-08-222-177A-137

Query Match	0.18;	Score 36;	DB 1;	Length 39;
Best Local Similarly	100.08;	Pred. No. 0.00084;		
Matches 36;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1 CACACACACACACACACACACACACACACA	36
Db	2 CACACACACACACACACACACACACACACA	37

RESULT 7

Mon Apr 3 08:24:17 2000

us-08-852-495c-1_copy_1_30000.rni

Page 4

```

US-08-222-177A-119          / Sequence 119, Application US/08222177A
/ Patent No. 5582979
/ GENERAL INFORMATION:
/ APPLICANT: Weber, James L.
/ TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
/ TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
/ NUMBER OF SEQUENCES: 460
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Demitt Ross & Stevens, S.C.
/ STREET: 8000 Excelsior Drive, Suite 401
/ CITY: Madison
/ STATE: Wisconsin
/ COUNTRY: USA
/ ZIP: 53717-1914
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/222,177A
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/341,562
/ FILING DATE: 21-APR-1989
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Sara, Charles S.
/ REGISTRATION NUMBER: 30,492
/ REFERENCE/DOCKET NUMBER: 09865, 601
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (608) 831-2100
/ TELEFAX: (608) 831-2106
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 119:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 40 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ IMMEDIATE SOURCE:
/ CLONE: mfd375
US-08-222-177A-119
Query Match      0.1%; Score 36; DB 1; Length 40;
Best Local Similarity 100.0%; Fred. No. 0.00084;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 CACACACACACACACACACACACACACACACACACA 36
        |||||||
Db       2 CACACACACACACACACACACACACACACACACA 37
RESULT      8
US-08-222-177A-400
/ Sequence 400, Application US/08222177A
/ Patent No. 5582979
/ GENERAL INFORMATION:
/ APPLICANT: Weber, James L.
/ TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
/ TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
/ NUMBER OF SEQUENCES: 460
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Demitt Ross & Stevens, S.C.
/ STREET: 8000 Excelsior Drive, Suite 401
/ CITY: Madison
/ STATE: Wisconsin
/ COUNTRY: USA
/ ZIP: 53717-1914
/ COMPUTER READABLE FORM:

```

```

MEDTUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865,601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX: (608) 831-2106
INFORMATION FOR SEQ ID NO: 400:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-222-177A-400

Query Match          0.1%; Score 36; DB I: Length 40;
Best Local Similarity 100.0%; Pred. No. 0.00084;
Matches      36; Conservative    0; Mismatches     0; Indels      0; Gaps      0;

QY       1 CACACACACACACACACACACACACACACACACACA 36
Db        2 CACACACACACACACACACACACACACACACACACA 37
|||||
|||||

RESULT 9
US-08-222-177A-403
Sequence 403. Application US/08222177A
Patent No. 5582879
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dc-dc)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Demilt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865,601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106

```


Mon Apr 3 08:24:17 2000

us-08-852-495c-1_copy_1_30000.mri

Page 8

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 03:15:46 ; Search time 8180.08 seconds
(without alignments)
13847.043 Million cell updates/sec

Title: US-08-852-495c-1_COPY_1_30000
Perfect score: 30000
Sequence: 1 CACACACACACACACACACA.....AGCCATACAGAGTGGCCCC 30000

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : ESR:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gp_est1:*
21: gp_est2:*
22: gp_est3:*
23: gp_est4:*
24: gp_est5:*
25: gp_est6:*
26: gp_est7:*
27: gp_est8:*
28: gp_est9:*
29: gp_est10:*
30: gp_est11:*
31: gp_est12:*
32: gp_est13:*
33: gp_est14:*
34: gp_est15:*
35: gp_est16:*
36: gp_est17:*
37: gp_est18:*
38: gp_est19:*
39: gp_est20:*
40: gp_est21:*
41: gp_est22:*
42: gp_est23:*
43: gp_est24:*

44: gp_est25:*
45: gp_est26:*
46: gp_est27:*
47: gp_est28:*
48: gp_est29:*
49: gp_est30:*
50: gp_est31:*
51: gp_est32:*
52: gp_est33:*
53: gp_est34:*
54: gp_est35:*
55: gp_est36:*
56: gp_est37:*
57: gp_est38:*
58: gp_est39:*
59: gp_est40:*
60: gp_est41:*
61: gp_est42:*
62: gp_est43:*
63: gp_est44:*
64: gp_est45:*
65: gp_est46:*
66: gp_est47:*
67: gp_est48:*
68: gp_est49:*
69: gp_est50:*
70: gp_est51:*
71: gp_est52:*
72: gp_est53:*
73: gp_est54:*
74: gp_est55:*
75: gp_est56:*
76: gp_est57:*
77: gp_est58:*
78: gp_est59:*
79: gp_est60:*
80: gp_est61:*
81: gp_est62:*
82: gp_est63:*
83: gp_est64:*
84: gp_est65:*
85: gp_est66:*
86: gp_est67:*
87: gp_est68:*
88: gp_est69:*
89: gp_est70:*
90: gp_est71:*
91: gp_est72:*
92: gp_est73:*
93: gp_est74:*
94: gp_est75:*
95: gp_est76:*
96: gp_est77:*
97: gp_est78:*
98: gp_est79:*
99: gp_est80:*
100: gp_est81:*
101: gp_est82:*
102: gp_est83:*
103: gp_est84:*
104: gp_est85:*
105: gp_est86:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	80	0.3	103	21	T77382 yd72h12.r1

Mon Apr 3 08:24:18 2000

us-08-852-495c-1_copy-1_30000.rst

Page 2

C	2	73	0.2	95	21	R08388
C	3	39	0.1	89	34	AA500598
C	4	38	0.1	95	20	T52775
C	5	38	0.1	101	90	AO006554
C	6	37	0.1	54	40	AA935225
C	7	37	0.1	58	68	HSM010004
C	8	37	0.1	80	50	E24490
C	9	37	0.1	89	36	AA598741
C	10	37	0.1	102	81	B03983
C	11	36	0.1	58	81	B03985
C	12	36	0.1	80	44	AI28205
C	13	36	0.1	84	50	F34634
C	14	36	0.1	89	28	AA078242
C	15	36	0.1	89	36	AA585741
C	16	36	0.1	97	28	AA077989
C	17	36	0.1	99	29	AA144438
C	18	36	0.1	101	63	AI964456
C	19	36	0.1	101	81	B15605
C	20	36	0.1	102	81	B47960
C	21	36	0.1	101	104	AA058286
C	22	35	0.1	103	38	AA796332
C	23	35	0.1	77	35	AA560791
C	24	35	0.1	85	81	B02822
C	25	35	0.1	86	38	AA745546
C	26	35	0.1	86	38	AA745546
C	27	35	0.1	91	50	E29274
C	28	34	0.1	47	68	HSM004698
C	29	34	0.1	78	50	E29721
C	30	34	0.1	83	40	AA988445
C	31	34	0.1	90	37	AA708008
C	32	34	0.1	94	43	AI203131
C	33	34	0.1	103	74	AM176421
C	34	34	0.1	57	46	AI434770
C	35	33	0.1	69	50	AI687854
C	36	33	0.1	83	28	AI171712
C	37	33	0.1	89	48	AI559683
C	38	33	0.1	91	48	AI453508
C	39	33	0.1	103	30	AA216932
C	40	32	0.1	55	36	AA601314
C	41	32	0.1	54	36	AA053911
C	42	32	0.1	58	43	AI120215
C	43	32	0.1	72	62	AI914922
C	44	32	0.1	73	70	AM152336
C	45	32	0.1	74	28	AA078709
C	2	73	0.2	95	21	R08388
C	3	39	0.1	89	34	AA500598
C	4	38	0.1	95	20	T52775
C	5	38	0.1	101	90	AO006554
C	6	37	0.1	54	40	AA935225
C	7	37	0.1	58	68	HSM010004
C	8	37	0.1	80	50	E24490
C	9	37	0.1	89	36	AA598741
C	10	37	0.1	102	81	B03983
C	11	36	0.1	58	81	B03985
C	12	36	0.1	80	44	AI28205
C	13	36	0.1	84	50	F34634
C	14	36	0.1	89	28	AA078242
C	15	36	0.1	89	36	AA585741
C	16	36	0.1	97	28	AA077989
C	17	36	0.1	99	29	AA144438
C	18	36	0.1	101	63	AI964456
C	19	36	0.1	101	81	B15605
C	20	36	0.1	102	81	B47960
C	21	36	0.1	101	104	AA058286
C	22	35	0.1	103	38	AA796332
C	23	35	0.1	77	35	AA560791
C	24	35	0.1	85	81	B02822
C	25	35	0.1	86	38	AA745546
C	26	35	0.1	86	38	AA745546
C	27	35	0.1	91	50	E29274
C	28	34	0.1	47	68	HSM004698
C	29	34	0.1	78	50	E29721
C	30	34	0.1			

RESULT	1				
LOCUS	T77382/c				
DEFINITION					
VERSION	T77382	103	bp	nrna	EST
KEYWORDS	T77382				
SOURCE	T77382.1	GI:694585			
ORGANISM	human.				
REFERENCE					
AUTHORS	1 (bases 1 to 103)				
	Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,				
	Holman,M., Hulcman,M., Kucaba,T., Le,M., Lennon,G., Merra,M.,				
	Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,				
	Trevaaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and				
	Wilson,R.				
TITLE	The WashU-Merck EST Project				
JOURNAL	Unpublished (1995)				
COMMENT	Other_ESTs: yd72h12.s1				
	Contact: Wilson RK				
	Washington University School of Medicine				

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel.: 314 286 1800
Fax: 314 286 1810
Email: estewatson.vustl.edu
Insert Size: 943
Source: IMAGE Consortium, LNL This clone is available royalty-free
through LNL: contact the IMAGE Consortium (info@image.lnl.gov)
for further information. Putative full length read
Insert Length: 943 Std Error: 0.00
Seq primer: M13R1
High quality sequence stop: 109.

FEATURES

```

/organism="Homo sapiens"
/db_xref="GDB:469448"
/bi_xref="taxon:9606"
/clone="IMAGE:113831"
/clone_id="Soares fetal liver spleen mFLs"
/sex="male"
/dev_stage="20 week post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/notes="organ: Liver and Spleen Vector: pRT3p (Pharmacia)
1st strand cDNA was primed with a Pac I, Sma2, Eco RI,
13 AACGCGAGAGATGATTAAGGTCCTTTTCTTTTCTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), sites of the modified pRT3 vector. Library
went through one round of normalisation. Library
constructed by Bento Soares and M.Fatima Bonaldo."

```

BASE COUNT
ORIGIN

	0.3%: Score 80: DB 21: Length 103:	
	Best Local Similarity 100.0%: Pred. No. 6,7e-16:	
Matches	80: Conservative	0: Mismatches 0: Indels 0: Gaps 0:
QY	6077 TAGAGTCTACCTCTACGCAACCTCCCTCTATATTCGAAGATTCCTGCTCA	6136
Db	103 TAGAGTCTACCTCTACGCAACCTCCCTCTATATTCGAAGATTCCTGCTCA	44
QY	6137 GCCTCCGAGTAGCTGGAC	6156
Db	43 GCCTCCGAGTAGCTGGAC	24

RESULT 2

LOCUS R08388

DEFINITION yf9403_r1 Soares fetal spleen INFRS Homo sapiens cDNA clone IMAGE127301 5' similar to gp:M60750_cds1 HISTONE H2B (HUMAN);, mRNA sequence.

ACCESSION R08388

VERSION R08388.1

KEYWORDS GI:760311

EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 95)

AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Parsons, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Meara, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trivaskar, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

source
1. 95
/organism="Homo sapiens"
/db_xref="GDB:50414"
/db_xref="taxon:9606"
/clone="IMAGE:6787"
/clone_lib="Stratagene ovary (#937217)"
/sex="female"
/dev_stage="49 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: ovary; Vector: Bluescript SK; Site_1: EcoRI;
Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt.
Total ovary tissue, normal, caucasian. Average insert
size: 0.8 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5'
GATTCGCGACGAG 3' -3' adaptor sequence: 5'
CTCAGATTTTCTTTTCTTTT 3' "

BASE COUNT 19 a 32 c 20 g 24 t

ORIGIN

Query Match 0.1%; Score 38; DB 20; Length 95;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13817 CAGTGCACATCTGGCTCAGTCACTCTGCTCC 13854
|||||
Db 27 CAGTGCACATCTGGCTCAGTCACTCTGCTCC 64

RESULT 5
AQ006534/c 101 bp DNA GSS 26-JUN-1998
LOCUS AQ006534
DEFINITION genomic survey sequence.
ACCESSION AQ006534
VERSION AQ006534.1 GI:3081351
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 101)
Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K.,
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Kim, U.-J.,
Shizuya, H., Simon, M., and Venter, J.C.
Use of a human BAC End Sequence Database for Sequence-Ready Map
Building
Unpublished (1997)
Other_GSS: CIT-HSP-2054B16.TR.1 CIT-HSP-2054B16.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/Bac_end_search/Bac_end_search.html
Seq primer: M13-21;
Class: BAC ends.

JOURNAL
TITLE
AUTHORS
REFERENCE
COMMENT

FEATURES

location/Qualifiers

1..101
/organism="Homo sapiens"
/db_xref="GDB:7057963"
/db_xref="taxon:9606"
/clone="2054B16"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBeloBAC11; site_1: HindIII; site_2:
HindIII"
BASE COUNT 25 a 12 c 24 g 40 t

ORIGIN

Query Match 0.1%; Score 38; DB 90; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACACACACACACACACACACACACACACACAA 38
|||||
Db 98 CACACACACACACACACACACACACACACACAAA 61

RESULT 6
AA935225/c 54 bp mRNA EST 07-JUL-1998
LOCUS AA935225
DEFINITION o06b11.s1 NCI-CGAP_G4 Homo sapiens cDNA clone IMAGE:1571133 3'
Similar to gb:144812 RERINOBLASTOMA-ASSOCIATED PROTEIN-LIKE 107 KD
HOMOLOG (HUMAN); contains LTR2, t2 LTR2 repetitive element ;, mRNA
sequence.
AA935225.1 GI:3092382
VERSION AA935225.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 54)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Jan 14, 1998 this sequence version replaced g1:1797423.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium LTRU at:
www.bio.lnl.gov/bbrp/image/image.html

JOURNAL
TITLE
AUTHORS
REFERENCE
COMMENT

FEATURES

location/Qualifiers

1..54
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1571133"
/clone_lib="NCI-CGAP_G4"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."
BASE COUNT 8 a 12 c 17 g 17 t

ORIGIN

Query Match 0.1%; Score 37; DB 40; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24632 CCATGCACTCGAGCTGGGACACAGCAAACTC 24658
|||||

Db 42 CCATTGCATCTCCAGCTGGCGCAACAGACAAACTC 6

RESULT 7
HSM010004/c
ID HSM010004 standard; RNA; EST; 58 BP.
XX
AC AL045154;
XX
SV AL045154.1
XX
DT 12-MAR-1999 (Rel. 59, Created)
DT 12-MAR-1999 (Rel. 59, last updated, Version 1)
XX
DE Homo sapiens mRNA; EST DKFZP434P234_r1 (from clone
DE DKFZP434P234)
XX
KM EST; expressed sequence tag.
XX
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria;
OC Primates; Catarrhini; Hominoidea; Homo.
XX
RN [1]
RP 1-58
RA Mambbutt R., Heubner D., Mewes W., Gassenhuber J., Wiemann S.;
RT Submitted (12-MAR-1999) to the EMBL/GenBank/DBJ databases.
RL MIPs, Am Klopferspitz 18a D-82152 Martinsried, GERMANY
XX
CC Clone from S. Wiemann, sequenced by AGOWA within the CDNA
CC sequencing consortium of the German Genome Project
CC s1 sequence also available at the RZPD in Berlin
CC This clone is available at the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
CC Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX
FH Key Location/Qualifiers
FH 1. 58
FT source
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT /clone="DKFZP434P234"
FT /clone_lib="434 (synonym: htes3). Vector pSPORT1; host
FT DH10B; sites NotI, SalI"
FT /dev_stage="adult"
FT /tissue_type="testis"
XX
SQ Sequence 58 BP; 13 A; 2 C; 21 G; 22 T; 0 other;

Query Match 0.1%; Score 37; DB 68; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 ACACACACACACACACACACACACACACACACAAA 38
Db 58 ACACACACACACACACACACACACACACACACAAA 22

RESULT 8
LOCUS F24490 80 bp mRNA EST 13-MAY-1999
DEFINITION HSPD10834 HM3 Homo sapiens cDNA clone s4000013A06, mRNA sequence.
ACCESSION F24490
VERSION F24490.1 GI:4810116
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 80)
Lanfranchi, G., Muraro, T., Caldara, F., Pacchioni, B., Pallavicini, A.,

TITLE
Pandofo, D., Toppo, S., Trevisan, S., Scarso, S., and Valle, G.
Identification of 4370 expressed sequence tags from a
3'-end-specific cDNA library of human skeletal muscle by DNA
sequencing and filter hybridization
Genome Res. 6 (1), 35-42 (1996)
96276048
On Jun 5, 1998 this sequence version replaced gi:3188834.
COMMENT
Contact: Valle G.
CIRI Biotechnology Centre
University of Padua
Via Trieste 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
http://grup.bio.unipd.it.
FEATURES
source
Location/Qualifiers
1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="s4000013A06"
/clone_lib="HM3"
/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/note="Vector: pCDNAII (Invitrogen); Site_1: BstXI;
Site_2: NotI. This library was constructed by G.
Lanfranchi. This library is not subtracted nor normalized.
The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-biotin-AACCGGCTCGAGCGCGCGCGCTTTTCTTTTCTTTT-3'). The
35 cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adapters, NotI digested and
directionally cloned into BstXI-NotI cut pCDNAII vector."

BASE COUNT
ORIGIN
18 a 21 c 19 g 22 t

Query Match 0.1%; Score 37; DB 50; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.087;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23005 GGGCGAGTGGCTCATGCTGTAAATCCACACTTTG 23041
Db 37 GGGCGAGTGGCTCATGCTGTAAATCCACACTTTG 73

RESULT 9
LOCUS AA598741/c
DEFINITION
ACCESSION AA598741
VERSION AA598741.1 GI:2432413
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 89)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellenberg, R., Steptoe, M., Tan, F.,
Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the

VERSION AI262095.1 GI:3870298
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 80)
AUTHORS NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 14, 1998 this sequence version replaced gi:1798653.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/BLNI at:
www-bio.1nl.gov/dbtrp/image/image.html

FEATURES
source
Seq primer: -40RP from Glibco.
Location/Qualifiers
1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:2028200"
/clone_lib="NCI_CGAP_K1d11"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7AD-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_K1d3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(clonids 132376-132391, 145607-145675, and
150052-150285). Subtraction by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 13 a 25 c 24 g 18 t

ORIGIN

Query Match 0.1%; Score 36; DB 44; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6248 CCCAGCTGAGTGCAGTGCACGATCTGGCTCAC 6283
|||||
Db 24 CCCAGCTGAGTGCAGTGCACGATCTGGCTCAC 59

RESULT 13
F34634 84 bp mRNA EST 13-MAY-1999
LOCUS HSPD29754 HM3 Homo sapiens cDNA clone sh3-000002-1/F06, mRNA
DEFINITION sequence.
ACCESSION F34634
VERSION F34634.1 GI:4820260
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 84)
AUTHORS Lanfranchi, G., Muraro, T., Caldara, F., Paschioni, B., Pallavicini, A.,
Pandofo, D., Toppo, S., Trevisan, S., Scarso, S. and Valle, G.
TITLE Identification of 4370 expressed sequence tags from a
3'-end-specific cDNA library of human skeletal muscle by DNA
sequencing and filter hybridization

JOURNAL Genome Res. 6 (1), 35-42 (1996)
MEDLINE 96276048
COMMENT On Jun 5, 1998 this sequence version replaced gi:3188981.
Contact: Valle G.
CIRI Biotechnology Centre
University of Padua
Via Trieste 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
<http://grup.bio.unipd.it>.

FEATURES
source
Location/Qualifiers
1..84
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="sh3-000002-1/F06"
/clone_lib="HM3"
/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/note="Vector: pCDM11 (Invitrogen); Site_1: BstXI;
Site_2: NotI; The library was constructed by G.
Lanfranchi. This library is not subtracted nor normalized.
The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-biotin-AACCGCGCTGAGCGCCGCTTTTCTTTTCTTTT-3'). The
ds cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adapters, NotI digested and
directionally cloned into BstXI-NotI cut pCDM11 vector."

BASE COUNT 18 a 17 c 23 g 25 t

ORIGIN

Query Match 0.1%; Score 36; DB 50; Length 84;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 23006 GGGCGAGTGGCTCATGCTGATCCAGCACTTGG 23041
|||||
Db 42 GGGCGAGTGGCTCATGCTGATCCAGCACTTGG 77

RESULT 14
AA078242/c 89 bp mRNA EST 24-SEP-1999
LOCUS 7H16G05 Chromosome 7 Hela cDNA library Homo sapiens cDNA clone
DEFINITION 7H16G05, mRNA sequence.
ACCESSION AA078242
VERSION AA078242.1 GI:1837716
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 89)
AUTHORS Touchman, J.W., Bouffard, G.G., Weintraub, L.A., Idol, J.R., Wang, L.,
Robbins, C.M., Nussbaum, J.C., Lovett, M. and Green, E.D.
TITLE 2006 expressed-sequence tags derived from human chromosome
7-enriched cDNA libraries
JOURNAL Genome Res. 7 (3), 281-292 (1997)
MEDLINE 97228905
COMMENT On Sep 12, 1996 this sequence version replaced gi:1292359.
Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892
Tel: 3014020201
Fax: 3014024735
Email: egreen@hgrl.nih.gov
Plate: 16 row: G column: 05
Seq primer: -21M13 (ABI).
Location/Qualifiers
1..89
/organism="Homo sapiens"

```

BASE COUNT
ORIGIN
0
a
c
45
g
44
t

```

```

Query Match          0.1%; Score 36; DB 28; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches    36; Conservative   0; Mismatches    0; Indels    0; Gaps    0;

```

RESULT	15
AA598741	
LOCUS	AA598741
DEFINITION	89 bp mRNA EST 19-DEC-1997
ACCESSION	aa69g2.1
VERSION	Stratagene lung carcinoma 93718 Homo sapiens CDNA clone
KEYWORDS	IMAGE:950278 3' similar to contains Alu repetitive element;; mRNA sequence.
SOURCE	AA598741.1 GI:2432413
ORGANISM	human.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
Eutheria; Primates; Carnivora; Homiidae; Homo.				
1 (bases 1 to 89)				
Hillier, L., Allen, M., Bowles, J., Dubouche, T., Giesel, G., Jost, S.,				
Kirman, D., Kucaba, T., Lacy, M., Le, N., Lemon, G., Marra, M.,				
Martin, J., Moore, B., Schellhammer, K., Stepien, M., Tan, F.,				
Theisinger, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.				
WashU-MCI human EST project				
Unpublished (1997)				
On May 18, 1995 this sequence version replaced gi:811447.				

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnlgov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham.

```

FEATURES
source
Location/Qualifiers
1. 89
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image:950278"
/clone_lib="Stratagene lung carcinoma 937218"
/tissue_type="lung carcinoma"
/cell_line="NCI-H69"
/dev_stage="cell line NCI-H69"
/lab_host="SOLR (Kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Small cell carcinoma cell line NCI-H69. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor
sequence: 5' GAGTTGGCAGCG 3' -3' adaptor sequence: 5'
CTGAGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT
13 a 25 g 29 t

```

```

ORIGIN
Query Match 0.1%; Score 36; DB 36; Length 89;
Best local Similarity 100.0%; Pred. No. 0.17;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12199 CCCAGGCTGGAGTGCATGCGCGGATCTGGCTCAC 12234
      |||||||||||||||||||||||||||||||
Db 28 CCCAGGCTGGAGTGCATGCGCGGATCTGGCTCAC 63

```

Search completed: March 29, 2000, 06:25:57
Job time: 11411 sec

```

Search completed: March 29, 2000, 06:25:57
Job time: 11411 sec

```

Mon Apr 3 08:24:18 2000

us-08-852-495c-1_copy_1_30000.rst

Page 9



•

•

•

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities. It emphasizes that this is crucial for ensuring transparency and accountability in the organization's operations.

2. The second part outlines the various methods and tools used to collect and analyze data. It mentions the use of both traditional and modern technologies to gather information from different sources.

3. The third part describes the process of interpreting the collected data and drawing meaningful conclusions. It highlights the need for a systematic approach to data analysis to avoid biases and errors.

4. The final part of the document provides recommendations for improving the data collection and analysis process. It suggests implementing standardized procedures and regularly updating the methods used to stay current with the latest practices.

OS Bos taurus.
PN MO9213102-A.
PD 06-AUG-1992.
PF 15-JAN-1992; U000340.
PR 15-JAN-1991; US-642342.
PA (GENM-) GENMARK.
PI Georges M, Massey JM;
DR WPI: 92-284684/34.
PT polymorphic bovine DNA markers - used in genetic identification,
PT genotyping, and selective breeding
PS Table 7; Page: 269; 517P; English.
CC The sequence is that of a bovine microsatellite sequence obt. by
CC screening a library of bovine MboI DNA fragments of between
CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MboI sites, the frequency of
CC (TC)₁₅ >9 microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for in vitro
CC amplification of the corresp. microsatellite (using the program
CC OPTIRIM). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved in the determinism of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SO Sequence 70 BP; 1 A; 1 C; 33 G; 35 T;

```

Query Match      0.2% Score 41; DB 1; Length 70;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches    41; Conservative   0; Mismatches   0; Indels
                                0; Gaps         0;

QY  7513 GGTGGTGGCTGTCGTGTCGTATGAGTGCGTGTGTCG 7553
      |||||||
Db   21 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 61

```

RESULT 3
 AC V19044/c
 ID V19044 standard; DNA; 40 BP.
 DT V19044.
 DT 28-JUL-1998 (first entry)
 DE Alu PCR primer 1.
 KW PCR; primer; amplification; Alu repeat sequence; vector;
 KW circular yeast artificial chromosome; YAC; ss.
 OS synthetic.
 OS Saccharomyces sp.
 PN MO9801573-A1.
 PD 15-JAN-1998.
 PE 09-JUL-1996; U11478.
 PR 09-JUL-1996; MO-U11478.
 PA (USSR) US DEPT HEALTH & HUMAN SERVICES.
 PI Koupinia NY, Iarionov VL, Perkins EL, Resnick MA;
 DR WPI; 98-110234/10.
 PT Preparation of yeast artificial chromosomes - by in vivo
 PT recombination using vector comprising yeast centromere, marker,
 PT yeast telomere and nucleic acid for recombination
 PS Example 1; Page 45; 117p; English.
 CC This is the nucleotide sequence for the PCR primer used in the
 CC amplification of the Alu repeat sequence, which is used to
 CC demonstrate the processes described in the invention. It involves
 CC the creation and use of circular yeast artificial chromosome (YAC)
 CC to selectively clone specific nucleic acids from a background of
 CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
 CC They can be used to rapidly isolate human DNA where only a part of the
 CC sequence of DNA is known. Using the methods large fragments of DNA can
 CC be easily cloned and analysed.
 CC Sequence 40 BP; 7 A; 12 C; 13 G; 8 T;

Query Match 0.28; Score 40; DB 1; Length 40;

```

Best Local Similarity 100.0%; Pred. No. 0.0042; Indels 0; Gaps 0;
Matches 40; Conservative 0; Mismatches 0;

Oy 8640 TCGGCTCCCAAGTCTGGATTACAGCGCTGAGCCACC 8679
      |||||||
Db 40 TCGGCTCCCAAGTCTGGATTACAGCGCTGAGCCACC 1

RESULT 4
T65714/C
ID T65714 standard; DNA; 39 BP.
AC T65714;
DT 17-JUN-1997 (first entry)
DE Repeat sequence from polymorphic marker clone Mcd12.
KM Polymorphism: repeat sequence; genetic marker; primer: amplification;
KW PCR; polymerase chain reaction; paternity; maternity; human; pedigree;
KW linkage analysis; genetic disease; animal; plant; breeding; locus;
KW hybridisation; chromosome; ds.
OS Homo sapiens.
PN US5582979-A.
PD 10-DEC-1996.
PR 21-APR-1989; 341562.
PR 21-APR-1989; US-341562.
PR 05-SEP-1991; US-754351.
PR 04-APR-1994; US-222177.
PA (MAR9-) MARSHFIELD CLINIC.
PI Weber JL;
DR WPI; 97-042299/04.
DE Detection of polymorphic genetic markers of the form
PT (dC-dA)n(dG-dT)n - using novel nucleic acid mols. as primers
PS (disclosure; Column 9-10; 186pp; English.
CC The invention relates to the isolation of polymorphic repeat sequences
CC having the sequence (dC-dA)n.(dG-dT)n which can be used as genetic
CC markers. Primers based on these sequences can be used to detect these
CC repeats, especially for use in e.g. paternity or maternity testing,
CC human genetic analysis such as linkage analysis of genetic disease,
CC commercial animal or plant breeding or pedigree analysis. Clones
CC containing the repeat sequences were isolated by hybridisation of
CC chromosome-specific phage libraries with a synthetic poly(dC-dA).(dG-dT)
CC probe. Over 100 repeat blocks were isolated. The inserts from the clones
CC were amplified by primers T65798-T66047. Those clones where the repeat
CC sequence has been determined are shown in T65704-797. This repeat
CC sequence is from the marker clone Mcd11 which contains the repeat
CC sequence having the formula (ACG)11AT(ACG)8A.
CC Sequence 39 BP; 20 A; 18 C; 0 G; 1 T;

Query Match 0.2%; Score 39; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 0.00088;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7522 TGTGTGTGTGTGTATGTATGTGTGTGTGTGTGTGT 7560
      |||||||
Db 39 TGTGTGTGTGTGTATGTATGTGTGTGTGTGTGTGT 1

RESULT 5
O34003
ID O34003 standard; DNA; 50 BP.
AC O34003;
DT 02-FEB-1993 (first entry)
DE Microsatellite sequence from clone TGLA400.
KM PCR; selection; primers: OPIPR1W; breeding; cattle; parentage;
KW genetic mapping; traits; amplification; ss.
OS Bos taurus.
PN WO8213102-A.
PD 06-AUG-1992.
PR 15-JUN-1992; U00340.
PR 15-JUN-1991; US-642342.
PA (GEN9-) GENMARK.
PI Georges M. Massey JM;
DR WPI; 92-264664/34.
DE Polymorphic bovine DNA markers - used in genetic identification,

```

PT Polymorphic bovine DNA markers - used in genetic identification,

gene mapping, and selective breeding
Table 7; page 331; 517pp; English.
The sequence is that of a bovine microsatellite sequence obt. by
screening a library of bovine MboI DNA fragments of between
250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
One out of 50 clones cross-hybridised. Assuming independent
distribution of microsatellites and MboI sites, the frequency of
(T6)_n >9 microsatellites in the bovine genome is estimated at >100,
000. The sequence information for ca. 250 such bovine microsatellites
is summarised in the specification and indexed herein (see below).
The sequences upstream and downstream of the microsatellite sequence
were used to generate the required PCR primers for in vitro
amplification of the corresp. microsatellite (using the program
OPIPRIM). The microsatellites may be used to identify individuals,
for parentage testing, and in the genetic mapping of economic trait
loci, or genes involved in the determinism of economically important
traits esp. in cattle, to allow selective breeding.
See also Q33501-34437.
Sequence 50 bp; 1 A; 2 C; 22 G; 25 T;

Query Match	0.2%	Score 38;	DB 1;	Length 50;
Best Local Similarity	100.0%	Pred.No. 0.0017;		
Matches 38; Conservative	0;	Mismatches 0;	Gaps 0	

```
Oy    7515 GAGGCTGTGTCTGTGTGTGTGTGTGAATGTCGACGTACTGT   7552
      |||||
Db     1 GTGGTGTGTAATACTGTGTGTGTGTGTGTGTGTGTGTGTGT   38
```

RESULT	6
Q34100	
ID	Q34100 standard; DNA; 57 BP.

DE Sequence of a microsatellite from clone TGLA58.
 KW PCR; selection; primers; OPTIPRM; breeding; cattle; parentage
 KW genetic mapping; traits; amplification; ss.

PD 06-AUG-1992.
PF 15-JAN-1992; U00340.
PR 15-JAN-1991; US-642342.

DR WPI: 92-284684/34.
PT Polymorphic bovine DNA markers - used in genetic identification
PT gene mapping, and selective breeding

CC by screening a library of bovine MBOI DNA fragments of between
CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites, the probability of finding a

CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required non-reducing form in vitro.

CC Optimprim). The microsatellites may be used to identify individuals/
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved in the determination of economically important
CC traits associated with the ability to allow collecting breeding

50 Sequence 57 BP; 1 A; 0 C; 27 G; 29 T;

	Query Match	0.28;	Score 38;	DB 1;	Length 57;
	Best Local Similarity	100.00;	Pred. No.	0.0017;	
	Matches 38;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	7513 GGTGCTGGTTCAGTGATCGTGCATTATGCTCCTGCTGCT	7550			

D_b

20 GTGTCTGTCGTCTGTCGTCTGTCATGCTGTCGT 57

RESULT	7
T65736/c	
ID	T65736 standard; DNA; 40 BP.

DE Repeat sequence from polymorphic marker clone Mfd34.
KM Polymorphism; repeat sequence; genetic marker; primer; amplification;
KW PCR; polymerase chain reaction; paternity; maternity; pedigree

OS	Homo sapiens.
PN	US5582979-A.
PD	10-DEC-1996.
DE	31-APR-1998. 341563

PR 05-SEP-1991; US-754351.
PR 04-APR-1994; US-222177.
PA (MARS-) MARSHFIELD CLINIC.
PI Weber, T.

PT (dc-da)n(dc-dt)n - using novel nucleic acid mols. as primers
 PT Disclosure; Column 9-10; 186pp; English.
 CC The invention relates to the isolation of polymorphic repeat

CC markers primers based on these sequences can be used to detect these
CC repeats, especially for use in e.g. paternity or maternity testing,
CC human genetic analysis such as linkage analysis of genetic disease,
CC commercial animal or plant breeding or pedigree analysis. Clones

CC chondrosarcoma-specific phage libraries with a synthetic poly(dA-cA)₁₂(dG-cG)₁₂ probe. Over 100 repeat blocks were isolated. The inserts from the clones
CC were amplified by primers T65798-T66047. Those clones where the repeat
CC sequence has been determined are shown in T65704-797. This repeat

CC	Sequence	40 BP;	20 A;	19 C;	0 G;	1 T;
CC	Sequence	40 BP;	20 A;	19 C;	0 G;	1 T;

Query Match	0.23;	Score 36;	DB 1;	Length 40;
Best Local Similarity	100.0%;	Pred. No. 0.008;		
Matches 36; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```

QY  7513 GTGCTGTGTGCTGTGTGTGTGTGTATGTGTGT 7548
      |||||
Db  36 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1

```

RESULT	8
Q27391	
ID	Q27391 standard; DNA; 35 BP.
RT	037301.

DE Inter-Alu specific primer Pbu33.
KW Polymerase chain reaction; PCR; repetitive element; ss
OS Synthetic.
PN W0823101-A

PF 24-JAN-1992; NL0018.
PR 25-JAN-1991; NL-000132.
PA (INGE-) INGENY BV.
PI titatortidogor NC v44c 1

PT Detection of genetic variation by 2-D electrophoresis of
PT fragments - and hybridisation with labelled probes, carried out
PT on fragments consisting of inter-repeat sequences generated by
PCR

PS Claim 6; Page 6; 31pp; English.

DT 14-SEP-1998 (first entry)
 DE Oligonucleotide sequence of the specification.
 KW Genotype: phenotype; molecular evolutionary engineering;
 KW functional biopolymer; virus; ss.
 OS Synthetic
 PN MO9816636-A1.
 PD 23-APR-1998.
 PR 17-OCT-1998; JP-274855.
 PA (MITU) MITSUBISHI CHEM CORP.
 PI Fusimi Y, Miyamoto E, Nemoto N, Yanagawa H;
 DR WPI; 98-261039/23.
 PT Virus containing nucleic acid and protein sections - for use in
 PT modification and creation of functional bio:polymers such as
 PT enzymes, antibodies and ribozyme(s)
 PS Example 1: Page 44; 68pp; Japanese.
 CC The present sequence is used in the course of the invention. The
 CC specification describes a molecule for bringing together genotype with
 CC phenotype (in vitro virus). The molecule contains a nucleic acid
 CC fragment having a base sequence corresponding to a genotype, covalently
 CC bonded to a protein fragment containing a protein participating in
 CC phenotype expression, the 3'-end of the nucleic acid part being bonded
 CC to the C-terminus of the protein part via a puromycin moiety. The nucleic
 CC acid fragment preferably consists of RNA corresponding to the gene (free
 CC from a termination codon), a spacer (such as two-stranded DNA), a peptide
 CC adapter (such as a DNA-tRNA hybrid containing an anticodon corresponding
 CC to the terminator codon of the gene) and a puromycin-containing cap
 CC capable of binding to an amino acid residue. Translation of the virus is
 CC performed in a non-cellular (preferably ribosomal) system (e.g. E. coli
 CC ribosome), the protein synthesised by the translation attaching to the
 CC puromycin cap, resulting in the complete in vitro virus structure. The
 CC method is used in molecular evolutionary engineering to optimise function
 CC of a functional biopolymer such as an enzyme, antibody or ribozyme, or
 CC to generate new functionality. The virus can be used for the functional
 CC optimisation of nucleic acid or protein sequences.
 SQ Sequence 80 BP; 7 A; 10 C; 1 G; 62 T;

Query Match 0.2%; Score 33; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 0.064;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 4426 AATAATTTTTTTTTTTTTTTTTTTTTTTT 4458
 DB 6 AATAATTTTTTTTTTTTTTTTTTTTTTTT 38

Search completed: March 29, 2000, 17:55:02
 Job time: 30700 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 12:57:56 ; Search time 229.02 Seconds

(without alignments)
10455.780 Million cell updates/sec

Title: US-08-852-495c-1_COPY_25000_45000

Perfect score: 20001

Sequence: 1 AGTCGAGTACGACCTC.....CAAGTTAGACATTTT 20001

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 214294 segs, 59861574 residues

Word size: 0

Total number of hits satisfying chosen parameters: 349160

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database:

Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/lna/3A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/lna/3B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/lna/3C.COMB.seq:*
- 4: /cgn2_6/ptodata/1/lna/3D.COMB.seq:*
- 5: /cgn2_6/ptodata/1/lna/6.COMB.seq:*
- 6: /cgn2_6/ptodata/1/lna/PTU59.COMB.seq:*
- 7: /cgn2_6/ptodata/1/lna/backfilssl.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	0.2	56	1	US-08-222-177A-149 Sequence 149, App
2	43	0.2	54	2	US-08-469-802B-27 Sequence 27, App
3	43	0.2	54	3	US-08-267-803B-45 Sequence 45, App
4	39	0.2	39	1	US-08-222-177A-86 Sequence 86, App
5	39	0.2	39	2	US-08-632-673B-1 Sequence 1, App
6	38	0.2	60	3	US-08-454-557C-57 Sequence 57, App
7	38	0.2	60	4	US-08-340-426D-57 Sequence 57, App
8	38	0.2	60	4	US-08-450-673C-57 Sequence 57, App
9	38	0.2	60	6	PCT-US95-17111A-57 Sequence 57, App
10	38	0.2	66	2	US-08-469-802B-29 Sequence 29, App
11	38	0.2	66	3	US-08-267-803B-47 Sequence 47, App
12	38	0.2	76	4	US-08-454-557C-69 Sequence 69, App
13	38	0.2	76	4	US-08-340-426D-69 Sequence 69, App
14	38	0.2	76	4	US-08-450-673C-69 Sequence 69, App
15	38	0.2	76	6	PCT-US95-17111A-69 Sequence 69, App
16	36	0.2	40	1	US-08-222-177A-152 Sequence 152, App
17	35	0.2	45	1	US-08-222-177A-157 Sequence 157, App
18	35	0.2	85	3	US-08-454-557C-92 Sequence 92, App
19	35	0.2	85	4	US-08-340-426D-92 Sequence 92, App
20	35	0.2	85	4	US-08-450-673C-92 Sequence 92, App
21	35	0.2	85	6	PCT-US95-17111A-92 Sequence 92, App
22	31	0.2	80	2	US-07-920-281C-25 Sequence 25, App
23	31	0.2	90	2	US-08-677-944-1 Sequence 1, App
24	30	0.1	37	1	US-08-113-646A-44 Sequence 44, App
25	30	0.1	43	1	US-08-222-177A-370 Sequence 370, App
26	30	0.1	56	4	US-08-776-944-9 Sequence 9, App

27	30	0.1	70	4	US-08-776-944-12 Sequence 12, App
28	30	0.1	75	4	US-08-776-944-13 Sequence 13, App
29	30	0.1	90	2	US-08-677-944-2 Sequence 2, App
30	29	0.1	30	2	US-08-433-505-9 Sequence 9, App
31	29	0.1	40	1	US-08-222-177A-175 Sequence 175, App
32	29	0.1	40	4	US-08-771-624B-1 Sequence 1, App
33	29	0.1	40	4	US-08-771-624B-10 Sequence 10, App
34	29	0.1	40	4	US-08-440-209-4 Sequence 4, App
35	29	0.1	42	1	US-07-875-167-2 Sequence 2, App
36	29	0.1	42	1	US-08-222-177A-53 Sequence 53, App
37	29	0.1	42	1	US-08-287-164-2 Sequence 2, App
38	29	0.1	44	1	US-08-113-646A-40 Sequence 40, App
39	29	0.1	44	2	US-08-664-596B-9 Sequence 9, App
40	29	0.1	47	4	US-08-778-494B-114 Sequence 114, App
41	29	0.1	50	1	US-08-233-609-5 Sequence 5, App
42	29	0.1	50	1	US-08-381-572-20 Sequence 20, App
43	29	0.1	50	1	US-08-444-083-5 Sequence 5, App
44	29	0.1	50	1	US-08-286-304-5 Sequence 5, App
45	29	0.1	50	1	US-08-420-443-1 Sequence 1, App

ALIGNMENTS

RESULT 1
US-08-222-177A-149/C
Sequence 149, Application US/08222177A
Patent No. 5582979

GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dc-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:

INFORMATION FOR SEQ ID NO: 149:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mid33its
US-08-222-177A-149

Query Match 0.2%; Score 48; DB 1; Length 56;


```

: REFERENCE/DOCKET NUMBER: 0609.3840004
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540
: INFORMATION FOR SEQ ID NO: 57:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 60 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: both
:
US-08-450-673C-57

Query Match          0.2%: Score 38: DB 4: length 60:
Best Local Similarity 100.0%: Pred. No. 9.2e-05:
Matches 38: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 12897 GGTGAACCCGTGCTCTACTATAAAATACAAAATTACG 12934
|||||
Db 48 GGTGAACCCGTGCTCTACTATAAAATACAAAATTACG 11

RESULT 9
PCT-US95-17111A-57/c
: Sequence 57, Application PC/TUS9517111A
: GENERAL INFORMATION:
: APPLICANT: de la Monte, Suzanne
: TITLE OF INVENTION: Neutral Thread Protein Gene Expression and
: TITLE OF INVENTION: Detection of Alzheimer's Disease
: NUMBER OF SEQUENCES: 121
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
: STREET: 1100 New York Avenue, Suite 600
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005-3934
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/17111A
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/340,426
: FILING DATE: 14-NOV-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Ludwig, Steven R.
: REGISTRATION NUMBER: 36,203
: REFERENCE/DOCKET NUMBER: 0609.3840002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540
: INFORMATION FOR SEQ ID NO: 57:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 60 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: both
:
PCT-US95-17111A-57

Query Match          0.2%: Score 38: DB 6: length 60:
Best Local Similarity 100.0%: Pred. No. 9.2e-05:
Matches 38: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Oy 12897 GGTGAACCCGTGCTCTACTATAAAATACAAAATTACG 12934
|||||
Db 48 GGTGAACCCGTGCTCTACTATAAAATACAAAATTACG 11

```

[illegible]

```

1 CITY: Minneapolis
2 STATE: MN
3 COUNTRY: USA
4 ZIP: 55458-1415
5
6 COMPUTER READABLE FORM:
7 MEDIUM TYPE: Floppy disk
8 COMPUTER: IBM PC compatible
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10 SOFTWARE: Patentin Release #1.0, Version #1.25
11
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER: US/08/267,803B
14 FILING DATE: 28-JUN-1994
15 CLASSIFICATION: 435
16
17 ATTORNEY/AGENT INFORMATION:
18 NAME: McCommack, Myra H.
19 REGISTRATION NUMBER: 36,602
20 REFERENCE/DOCKET NUMBER: 110,00030120
21
22 TELECOMMUNICATION INFORMATION:
23 TELEPHONE: 612-305-1217
24 TELEFAX: 612-305-1228
25
26 INFORMATION FOR SEQ ID NO: 47:
27 SEQUENCE CHARACTERISTICS:
28 LENGTH: 66 base pairs
29 TYPE: nucleic acid
30 STRANDEDNESS: single
31 TOPOLOGY: linear
32
33 MOLECULE TYPE: DNA
34
35 US-08-267-803B-47
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
```

SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-69

Query Match 0.2%; Score 38; DB 3; Length 76;
Best Local Similarity 100.0%; Pred. No. 8.9e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12897 GGTGAACCCGTCTCTACTACAAATAATAGC 12934
|||||
DB 48 GGTGAACCCGTCTCTACTACAAATAATAGC 11

RESULT 13
US-08-340-426D-69/C
Sequence 69 Application US/08340426D
Patent No. 5948834
GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:

SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-69

Query Match 0.2%; Score 38; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 8.9e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12897 GGTGAACCCGTCTCTACTACAAATAATAGC 12934
|||||
DB 48 GGTGAACCCGTCTCTACTACAAATAATAGC 11

RESULT 14
US-08-450-673C-69/C
Sequence 69 Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:

SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-69

Query Match 0.2%; Score 38; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 8.9e-05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12897 GGTGAACCCGTCTCTACTACAAATAATAGC 12934
|||||
DB 48 GGTGAACCCGTCTCTACTACAAATAATAGC 11

RESULT 15
PCT-US95-17111A-69/C
Sequence 69 Application PCT/US9517111A
GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426

FILING DATE: 14-NOV-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Ludwig, Steven R.
 REGISTRATION NUMBER: 36,203
 REFERENCE/DOCKET NUMBER: 0609.3840002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 69:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 76 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: both
 PCT-US93-17111A-69

Query Match 0.2%; Score 38; DB 6; Length 76;
 Best Local Similarity 100.0%; Pred. No. 8.9e-05;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12897 GGCGAAACCTGTCTCTACTAAATAACAAATAATAGC 12934
 ||||||||||||||||||||||||||||||||||||
 DB 48 GGCGAAACCTGTCTCTACTAAATAACAAATAATAGC 11

Search completed: March 29, 2000, 17:48:08
 Job time: 50350 sec



4

.

-

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 06:25:57 ; Search time 5477.05 Seconds
(without alignments)
13787.907 Million cell updates/sec

Title: US-08-852-495c-1_COPY_25000_45000
Perfect score: 20001
Sequence: 1 AGTGCAGTAGCTGCACCTC.....CAAGTTAGACATTTT 20001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : EST:
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est33:*
53: gb_est34:*
54: gb_est35:*
55: gb_est36:*
56: gb_est37:*
57: gb_est38:*
58: gb_est39:*
59: gb_est40:*
60: gb_est41:*
61: gb_est42:*
62: gb_est43:*
63: gb_est44:*
64: gb_est45:*
65: gb_est46:*
66: gb_est47:*
67: gb_est48:*
68: gb_est49:*
69: gb_est50:*
70: gb_est51:*
71: gb_est52:*
72: gb_est53:*
73: gb_est54:*
74: gb_est55:*
75: gb_est56:*
76: gb_est57:*
77: gb_est58:*
78: gb_est59:*
79: gb_est60:*
80: gb_est61:*
81: gb_est62:*
82: gb_est63:*
83: gb_est64:*
84: gb_est65:*
85: gb_est66:*
86: gb_est67:*
87: gb_est68:*
88: gb_est69:*
89: gb_est70:*
90: gb_est71:*
91: gb_est72:*
92: gb_est73:*
93: gb_est74:*
94: gb_est75:*
95: gb_est76:*
96: gb_est77:*
97: gb_est78:*
98: gb_est79:*
99: gb_est80:*
100: gb_est81:*
101: gb_est82:*
102: gb_est83:*
103: gb_est84:*
104: gb_est85:*
105: gb_est86:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	51	0.3	70	25	N84707

N84707 J0579F Huma

2	47	0.2	59	28	AA082835	AA082835 zn21g12.s
3	47	0.2	59	28	AA082835	AA082835 zn21g12.s
4	44	0.2	103	104	AA053524	AA053524 RPT-11-3
5	43	0.2	102	81	AA0808	AA0808 RPT-11-4N6
6	40	0.2	68	28	AA4637	AA4637 HUSUPY036
7	40	0.2	68	28	AA4637	AA4637 HUSUPY036
8	39	0.2	100	26	AA7414	AA7414 ZD0305.11
9	38	0.2	101	35	AA53145	AA53145 n07g04.s
10	38	0.2	101	35	AA53145	AA53145 n07g04.s
11	38	0.2	102	100	AA019270	AA019270 RPT-11-98
12	38	0.2	102	100	AA019270	AA019270 RPT-11-98
13	38	0.2	105	36	AA614379	AA614379 n046c03.s
14	37	0.2	105	36	AA614379	AA614379 n046c03.s
15	37	0.2	72	45	AA134913	AA134913 t051g08.x
16	37	0.2	105	28	AA078003	AA078003 t051g08.x
17	36	0.2	105	61	AA183283	AA183283 t051g08.x
18	36	0.2	102	81	AA280198	AA280198 t051g08.x
19	35	0.2	102	81	AA280198	AA280198 t051g08.x
20	35	0.2	95	35	AA578401	AA578401 n153c01.s
21	35	0.2	96	38	AA747857	AA747857 n031f09.s
22	34	0.2	54	36	AA601314	AA601314 n015f06.s
23	34	0.2	90	37	AA708028	AA708028 t051g08.x
24	34	0.2	94	43	AA120313	AA120313 t051g08.x
25	34	0.2	99	48	AA158327	AA158327 t051g08.x
26	34	0.2	101	35	AA235077	AA235077 t051g08.x
27	34	0.2	101	35	AA235077	AA235077 t051g08.x
28	34	0.2	101	35	AA583657	AA583657 n058f10.s
29	34	0.2	103	74	AA176421	AA176421 t051g08.x
30	34	0.2	103	74	AA176421	AA176421 t051g08.x
31	33	0.2	57	100	AA0276193	AA0276193 t051g08.x
32	33	0.2	57	100	AA0276193	AA0276193 t051g08.x
33	33	0.2	69	50	AA168764	AA168764 t051g08.x
34	33	0.2	78	28	AA117122	AA117122 t051g08.x
35	33	0.2	83	28	AA117122	AA117122 t051g08.x
36	33	0.2	85	44	AA123453	AA123453 t051g08.x
37	33	0.2	85	44	AA123453	AA123453 t051g08.x
38	33	0.2	89	62	AA159863	AA159863 t051g08.x
39	33	0.2	89	62	AA159863	AA159863 t051g08.x
40	33	0.2	90	44	AA131403	AA131403 t051g08.x
41	33	0.2	91	46	AA153508	AA153508 t051g08.x
42	33	0.2	95	35	AA457423	AA457423 t051g08.x
43	33	0.2	95	35	AA457423	AA457423 t051g08.x
44	33	0.2	99	34	AA120282	AA120282 t051g08.x
45	33	0.2	100	35	AA564832	AA564832 t051g08.x

ALIGNMENTS

RESULT 1	70 bp	MRNA	EST	01-APR-1996
LOCUS	70 bp	MRNA	EST	01-APR-1996
DEFINITION	Human fetal heart, lambda ZAP Express Homo sapiens cDNA			
ACCESSION	J05795.5			
VERSION	Similar to REPETITIVE ELEMENT ALU, mRNA sequence.			
KEYWORDS	EST.			
SOURCE	Human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
AUTHORS	Eutheria; Primates; Catarrhini; Homiidae; Homo.			
JOURNAL	Liew, C.C.			
TITLE	CDNAs from fetal heart (1996)			
COMMENT	Unpublished (1996)			
	Contact: Liew CC			
	Department of Laboratory Medicine and Pathobiology			
	University of Toronto			
	Banting Institute, 100 College St., Toronto, Ontario, M5G1L5			
	Tel: 4169788758			
	Fax: 4169788550			
	Email: liewc@utoronto.ca			
	Seq primer: GAAATTACCTCTCAAGG.			

FEATURES

1..70
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="10579"
 /clone_lib="human fetal heart, lambda ZAP Express"
 /lab_host="E. coli XL1-Blue"
 /note="Vector: lambda ZAP Express; Site-1: EcoRI, Site-2: XhoI; mRNA was purified from human fetal hearts (8-10 weeks). cDNA was synthesized using a XhoI-Oligo dT adaptor-primer. EcoRI adaptors were ligated, followed by digestion with XhoI, for directional cloning into predigested lambda ZAP Express."

BASE COUNT
 24 a 18 c 15 g 13 t

Query Match 0.3%; Score 51; DB 25; Length 70;
 Best Local Similarity 100.0%; Pred. No. 8.4e-07;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13907 TAGAGTGGGTTTCCATGTCAGGATGTCGATCTGACT 13957

Db 57 TAGAGTGGGTTTCCATGTCAGGATGTCGATCTGACT 7

RESULT 2

LOCUS	AA082835	59 bp	MRNA	EST	21-OCT-1996
DEFINITION	zn21g12.s1 stragene neuroepithelium NT2RAMI 937234 Homo sapiens				
ACCESSION	AA082835				
VERSION	AA082835.1	GI:1624910			
KEYWORDS	EST.				
SOURCE	Human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
AUTHORS	Eutheria; Primates; Catarrhini; Homiidae; Homo.				
	1 (bases 1 to 59)				
	Chissoe, S., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,				
	Hawkins, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,				
	Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,				
	Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.,				
	and Marra, M.				
	Generation and analysis of 280,000 human expressed sequence tags				
	Genome Res. 6 (9), 807-828 (1996)				
TITLE	On Sep. 12, 1996 this sequence version replaced gi:1404580.				
JOURNAL	Washington University School of Medicine				
MEDLINE	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108				
COMMENT	Tel: 314 286 1800				
	Fax: 314 286 1810				
	Email: estevan@wustl.edu				
	This clone is available royalty-free through LINT; contact the				
	IMAGE Consortium (info@image.llnl.gov) for further information.				
	Seq primer: -40M13 fwd, from Amersham.				
FEATURES	Location/Qualifiers				
source	1..59				
	/organism="Homo sapiens"				
	/db_xref="GDB:392650"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:548134"				
	/clone_lib="Stragene neuroepithelium NT2RAMI 937234"				
	/dev_stage="Ntera-2/RA+MI neuroepithelial cells"				
	/lab_host="SOLR (kanamycin resistant)"				
	/note="Vector: pBluescript SK-; Site-1: EcoRI; Site-2: XhoI; Cloned unidirectionally. Primer: Oligo dT, NT2				
	(Ntera-2/c1.D1) precursor cells induced with RetA				
	acid for 1 week, followed by 3 weeks in mitotic inhibitors				

(Replate #2). Average insert size: 1.1 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGACAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3' "

BASE COUNT 11 a 14 c 16 g 18 t

Query Match 0.2%; Score 47; DB 28; Length 59;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13914 GGGGTTTCACCATGTTAGCCAGATGCTCGATCTCGTACCTCT 13960

DB 13 GGGGTTTCACCATGTTAGCCAGATGCTCGATCTCGTACCTCT 59

RESULT 3

AI914923 72 bp mRNA EST 28-JUL-1999

LOCUS tr27e08.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2219558 3'

DEFINITION similar to contains Alu repetitive element; mRNA sequence.

AI914923 AI914923.1 GI:5634778

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 72)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

On May 1, 1997 this sequence version replaced GI:2059395.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bdrip/image/image.html

Seq primer: -40UP from Gibco.

Location/Qualifiers

1..72

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="2219558"

/clone_lib="NCI_CGAP_Ov23"

/tissue_type="tumor, 5 pooled (see description)"

/lab_host="DH10B"

/note="Organ: ovary; Vector: pCMV-SPCKT6; Site_1: Salt;

Site_2: Not; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.35 kb. Tumor types include: mixed

Mullerian tumor, papillary serous, clear cell, spindle

cell. All are primary tumors, metastasis positive. Life

Technologies catalog #: 11534-013"

BASE COUNT 8 a 11 c 15 g 38 t

ORIGIN

Query Match 0.2%; Score 47; DB 62; Length 72;

Best Local Similarity 100.0%; Pred. No. 1.8e-05;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18759 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 18805

DB 4 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 50

RESULT 4

AO535244

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

COMMENT

JOURLNL

source

1..103

/organism="Homo sapiens"

/db_xref="GDB:7621533"

/db_xref="taxon:9606"

/clone="RPCI-11-317H22"

/clone_lib="RPCI-11"

/sex="Male"

/cell_type="Lymphocytes"

/note="vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;

RPCI11 Human Male BAC Library"

BASE COUNT 31 a 27 c 27 g 18 t

ORIGIN

Query Match 0.2%; Score 44; DB 104; Length 103;

Best Local Similarity 100.0%; Pred. No. 0.00015;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10109 TTGAGTTCGAGAGTTCGAGACCAAGCTGGCCAACTGGTGAAC 10152

DB 35 TTGAGTTCGAGAGTTCGAGACCAAGCTGGCCAACTGGTGAAC 78

RESULT 5

B48088

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 102)

Adams, M.D., Bounslay, S.D., Field, C.E., Bass, S., Linher, K.,

Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and

TITLE Venter, J.C.
JOURNAL Unpublished (1997)
COMMENT Other GSS: RPCI11-4N6-TV
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/human/bac_end_search/bac_end_search.html
Seq primer: sp6
Class: BAC ends.

FEATURES
SOURCE Location/Qualifiers

1. 102
/organism="Homo sapiens"
/db_xref="GDB:7501469"
/db_xref="taxon:9606"
/clone="RPCI-11-4N6"
/clone_1lb="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector PBAC3.6, Site 1: EcoRI, Site 2: EcoRI, RPCI11 Human Male BAC library"
BASE COUNT 32 a 24 c 25 g 21 t
ORIGIN

Query Match 0.2%; Score 43; DB 81; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12887 TGGCCACATGCTGAACCCCTCTCTACTATAAAATACAAA 12929
Db 24 TGGCCACATGCTGAACCCCTCTCTACTATAAAATACAAA 66

RESULT 6
LOCUS D44637/c 68 bp mRNA EST 20-PEB-1998
DEFINITION HNSUPY036 Human brain cDNA Homo sapiens cDNA clone 033-00-1, mRNA sequence.
ACCESSION D44637
VERSION D44637.1 GI:1572112
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 68)
Hadao, S., Ishida, Y., Tomiyasu, H., Yamamoto, K., Bates, G. P. and Trede, J.

TITLE Transcript map of the human chromosome 4p16.3 consisting of 627 DNA clones derived from 1 Mb of the Huntington's disease locus
JOURNAL DNA Res. 3 (4), 239-255 (1996)
MEDLINE 97101646
COMMENT On May 5, 1995 this sequence version replaced g1:798369.
Contact: Shinji Hadao
Japan Science and Technology Corporation, Neurogenes Project, ICOMP Univ. of Tokai School of Med.
Bohsedai, Isehara, Kanagawa 259-1193, Japan
Tel: 81-463-91-5095
Fax: 81-463-91-4993
Email: shinji@eng.med.u-tokai.ac.jp.

FEATURES
SOURCE Location/Qualifiers
1. 68
/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="033-00-1"
/clone_1lb="Human brain cDNA"
/tissue_type="brain"
BASE COUNT 14 a 16 c 25 g 13 t
ORIGIN

Query Match 0.2%; Score 40; DB 28; Length 68;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8495 GCAATTCCTGCTCAGCCTCCGAGTAGCTGGGATTAC 8534
Db 68 GCAATTCCTGCTCAGCCTCCGAGTAGCTGGGATTAC 29

RESULT 7
LOCUS R67088 96 bp mRNA EST 30-MAY-1995
DEFINITION Y130H05.s1 Soares placenta NB2HP Homo sapiens cDNA clone IMAGE:140793 3', mRNA sequence.
ACCESSION R67088
VERSION R67088.1 GI:839726
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 96)
Hallier, L., Clark, N., Dubouque, T., Elliston, K., Hawkins, M., Holman, M., Huitema, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Ritzkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.
The Wasnu-Werck EST Project
Unpublished (1995)
On Apr 5, 1995 this sequence version replaced g1:760768.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1093
High quality sequence stops: 65
Source: IMAGE Consortium, LNLN.
This clone is available royalty-free through LNLN; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Insert length: 1093 Std Error: 0.00
Seq primer: Promega -21m13
High quality sequence stop: 65.

TITLE Unpublished (1995)
JOURNAL On Apr 5, 1995 this sequence version replaced g1:760768.
COMMENT

FEATURES

SOURCE

1. 96
/organism="Homo sapiens"
/db_xref="GDB:349405"
/db_xref="taxon:9606"
/clone="IMAGE:140793"
/clone_1lb="Soares placenta NB2HP"
/sex="Female"
/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Placenta; Vector: p773D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' AACTGGAGAAATTCGCGCCGAGAAATTTTATTTTATTTT 3'), double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Patricia Bonalao.

BASE COUNT 25 a 24 c 21 g 25 t
ORIGIN

[illegible]

```

LOCUS       AA555145/c                      101 bp      mRNA                     EST               12-AUG-1997
DEFINITION  n107g64.s1 NCI-CGAP_Prl1 Homo sapiens CDNA clone IMAGE:1029654.
            similar to contains Alu repetitive element;contains element TANI
            repetitive element ;, mRNA sequence.
ACCESSION   AA555145
VERSION     AA555145.1  GI:2325684
KEYWORDS    EST.

SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 101)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     On Sep 12, 1996 this sequence version replaced gi:1397814.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert.Strausberg@nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Rodrigo F.
            Chuagui, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Kitzman, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LMWL at:
            www-bio.llnl.gov/bdrrp/image/lmage.html

FEATURES             source          location
Seq primer: -40m13 fwd, RT from Amersham
High quality sequence stop: 84.
Location/Qualifiers
     1..101
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
         /clone="IMAGE:1029654"
         /clone_1lb="NCI-CGAP_Prl1"
         /sex="male"
         /tissue_type="normal prostatic epithelial cells"
         /lab_host="DH10B"
         /note="Organ: prostate; Vector: pAMP10; mRNA made from
         normal prostatic epithelial cells. cDNA made by oligo-dT
         priming. Non-directionally cloned. Size-selected on
         agarose gel, average insert size 600 bp. Library made by
         D. Kitzman, NIH."

BASE COUNT        66 a      14 c      10 g      11 t

ORIGIN
Query Match           0.2%; Score 39; DB 35; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.0074;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
cy 18759 TTTTTTTTTTTTTTTTTTTTTTGAGACAGACT 18797
|||||
Db 75 TTTTTTTTTTTTTTTTTTTTTTGAGACAGACT 37

RESULT  10
LOCUS    H67549                      98 bp      mRNA                     EST               11-DEC-1995
DEFINITION  yu68f10.s1 Weizmann Olfactory Epithelium Homo sapiens cDNA clone
            IMAGE:238987 3' similar to contains Alu repetitive element;contains
            MER22 repetitive element ;, mRNA sequence.
ACCESSION   H67549
VERSION     H67549.1  GI:1114431
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 98)

```

AUTHORS

Hallier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapella, B.,
 Haisoe, S., Dietrich, N., Dubuque, T., Favell, A., Gish, M.,
 Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,
 Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rittin, L.,
 Roeding, T., Schellenberg, K., Soares, M.B., Tan, F., Thery-Meg, J.,
 Trevasaki, E., Underwood, K., Wohlman, P., Waterston, R., Wilson, R.,
 and Maria, M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)
 97044478

JOURNAL
MEDLINE

On Oct 18, 1995 this sequence version replaced g1:1026289.

CONTACT: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 Insert Size: 174
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: Promega -21m13
 location/Qualifiers

FEATURES
source

1. 98
 /organism="Homo sapiens"
 /db_xref="GDB:3864371"
 /db_xref="taxon:9606"
 /clone_image="238987"
 /clone_lib="Weizmann Olfactory Epithelium"
 /sex="Female"
 /tissue_type="Olfactory epithelium"
 /dev_stage="35 year old"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="Organ: nose; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
 Oligo dt. Olfactory epithelium, normal. Average insert
 size: 0.8 kb; Uni-ZAP XR Vector. Library constructed by N.
 Walker, D. Lancet, Weizmann Institute of Science. -5'
 adaptor sequence: 5' GAATTCGGCGACG 3' -3' adaptor
 sequence: 5' CTCGACTTTTCTTTTCTTTT 3'."
 BASE COUNT 19 a 39 c 23 g 17 t
 ORIGIN

Query Match 0.2%; Score 38; DB 24; Length 98;
 Best Local Similarity 100.0%; Pred. No. 0.016;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8481 CCTCCAGTCAAGCAATCTCTGCTCAGCCGCC 8518
 DB 50 CCTCCAGTCAAGCAATCTCTGCTCAGCCGCC 87

RESULT 11
 A0319270 102 bp DNA GSS 06-MAY-1999
 LOCUS R0311-98B22-TV R031-11 Homo sapiens genomic clone R031-11-98B22,
 DEFINITION genomic survey sequence.
 ACCESSION A0319270
 VERSION A0319270.1 GI:4052235
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eukaryota; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Base, S., Liner, K., Golden, K.,
 Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P., and Venter, J.C.
 Title: Human BAC End Sequences for Sequence-Ready Map Building
 Journal: Unpublished (1998)
 Comment: Other-GSSs: R0311-98B22-TV
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics

The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbe@igrr.org
 Clones are derived from the human BAC library R031-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buitalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buitalo.edu/ordering) or from
 Research Genetics (info@resgen.com). BAC end search page:
 http://www.tigr.org/tdb/hungen/bac_end_search.html
 Seq primer: Sp6
 Class: BAC ends.
 location/Qualifiers

FEATURES
source

1. 102
 /organism="Homo sapiens"
 /db_xref="GDB:7537293"
 /db_xref="taxon:9606"
 /clone_image="R031-11-98B22"
 /clone_lib="R031-11"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
 R0311 Human Male BAC library"
 BASE COUNT 26 a 28 c 31 g 17 t
 ORIGIN

Query Match 0.2%; Score 38; DB 100; Length 102;
 Best Local Similarity 100.0%; Pred. No. 0.016;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17275 CCACCGGCTGGGCTCCCAAGTGTGGATTACAG 17312
 DB 91 CCACCGGCTGGGCTCCCAAGTGTGGATTACAG 54

RESULT 12
 AA614379 104 bp mRNA EST 06-OCT-1997
 LOCUS np46c03.31 NCI-CGAP Brl.1 Homo sapiens cDNA clone IMAGE:1129348 3'
 DEFINITION similar to contains Alu repetitive element; mRNA sequence.
 ACCESSION AA614379
 VERSION AA614379.1 GI:2465575
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eukaryota; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 Title: National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 On May 5, 1995 this sequence version replaced g1:798056.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bdnp/image/image.html

Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop 52.
 location/Qualifiers

FEATURES
source

1. 104
 /organism="Homo sapiens"
 /db_xref="taxon:9606"


```

/clone="IMAGE:1129348"
/clone_lib="NCI_CGAP_Brl.1"
/sex="female, pooled"
/tissue_type="breast"
/lab_host="DH10B"
/ncore="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker: 1st strand cDNA was prepared from pooled bulk
breast tumor tissue, and was then primed with a Not I -
o119o(dt) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is not normalized. (The normalized
version of this library is NCI_CGAP Br2.) Library was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      19 a      26 c      29 g      30 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 38; DB 36; Length 104;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Oy 8635 CCGCCTCGGCTCCCAAGTCTGGGATTACAGGCGTG 8672
|||||
Db 67 CCGCCTCGGCTCCCAAGTCTGGGATTACAGGCGTG 104

RESULT 13
LOCUS AI752407/c 105 bp mRNA EST 22-JUN-1999
DEFINITION cni5d10.v1 Normal Human Trabecular Bone Cells Homo sapiens CDNA
ACCESSION AI752407
VERSION AI752407.1 GI:5130671
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 105)
AUTHORS Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G.,
Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M.,
Robey,P.G., Hotchkiss,R.N. and Franccomano,C.A.
SGAP: The Skeletal Genome Anatomy Project
Unpublished (1997)
COMMENT On Dec 20, 1995 this sequence version replaced gi:1130959.
Contact: Libin Jia
Medical Genetics Branch
National Human Genome Research Institute
10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
Tel: 301-402-4877
Fax: 301-496-7157
Email: libin@helix.nih.gov
DNA Sequencing and analyses by National Institutes of Health
Intramural Sequencing Center (NISC).
Plate: 15 row: d column: 10
Seq primer: M13RP1 reverse primer (ABI).
FEATURES
Source
1..105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCBI_Cn15d10"
/clone_lib="Normal Human Trabecular Bone Cells"
/sex="female"
/tissue_type="Bone"
/cell_type="Trabecular Bone Cells"
/lab_host="SURE"
/note="Organ: Hip; Vector: pBluescript; Site_1: EcoRI"
BASE COUNT      34 a      22 c      26 g      22 t      1 others
ORIGIN

Query Match      0.2%; Score 38; DB 51; Length 105;

```

QY	8646	TCCCAAAGTCTGGGATTACAGGCGGTGAGCACCACGC	8683
Db	105	TCCCAAAGTCTGGGATTACAGGCGGTGAGCACCACGC	68
RESULT 14			
LOCUS	AI349123/c		
DEFINITION	tab1908.x2 NCI_CGAP_Lu25 Homo sapiens cDNA clone IMAGE:2047646 3', mRNA sequence.		
ACCESSION	AI349123		
VERSION	AI349123.1	GI:4086329	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 72)		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert_Strausberg@nih.gov CDNA Library Preparation: David B. Kitzman, Ph.D. DNA Library Arrayed by: I.M.A.G.E. Consortium, LNL DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www-bio.lnl.gov/bdrip/image/image.html		
FEATURES			
SOURCE	Seq primer: -40UP from GIBCO. Location/Qualifiers		
	1..72		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:2047646"		
	/clone_lib="NCI CGAP_Lu25"		
	/tissue_type="bronchioalveolar carcinoma"		
	/dev_stage="adult"		
	/lab_host="DH10B"		
	/note="Organ: Lung; Vector: pAMP1; mRNA made from lung carcinoma tissue, cDNA made by oligo-dT priming. directionally cloned. Size-selected on agarose gel, average insert size 500 bp. Primary library, non-amplified."		
BASE COUNT	40 a	8 c	17 g
ORIGIN			7 t
Query Match		0.24;	Score 37; DB 45; Length 72;
Best Local Similarity	100.04;	Pred. No. 0.042;	
Matches	37; Conservative	0; Mismatches	0; Indels
			0; Gaps
QY	18764	TTTTTTTTTTTTTTTTTTTTCAGACAGAGTCTC	18800
Db	62	TTTTTTTTTTTTTTTTTTTTCAGACAGAGTCTC	26
RESULT 15			
LOCUS	AA078003/c		
DEFINITION	7H12D08 Chromosome 7 HeLa cDNA Library Homo sapiens cDNA clone IMAGE:2047646 3', mRNA sequence.		
ACCESSION	AA078003		
VERSION	AA078003.1	GI:1837477	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 105)
Touchman, J.W., Bouffard, G.G., Weintraub, L.A., Idol, J.R., Wang, L.,
Robbins, C.M., Nussbaum, J.C., Lovett, M., and Green, E.D.
2006 expressed-sequence tags derived from human chromosome
7-enriched cDNA libraries
Genome Res. 7 (3), 281-292 (1997)
JOURNAL
MEDLINE
97228905
COMMENT
On Apr 14, 1993 this sequence version replaced gi:693433.
Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892
Tel.: 3014020201
Fax: 3014024735
Email: egreen@nhgri.nih.gov
Plate: 12 row: D column: 08
Seq primer: -21M13 (ABI).
Location/Qualifiers

FEATURES
source
1..105
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="7H12D08"
/clone_lib="Chromosome 7 HeLa cDNA Library"
/sex="female"
/cell_line="HeLa cell line; ATCC"
/lab_host="E. coli strain DH5 alpha"
/note="Vector: PAMP10; cDNA was generated from cytoplasmic
RNA using a mixture of random DNA hexamers and oligo(dT).
From this pool of cDNA, human chromosome 7-enriched cDNA
was isolated by direct cDNA selection using chromosome 7
genomic DNA (cosmids). The resulting direct-selected cDNA
was cloned into a plasmid vector using a non-directional
uracil DNA glycosylase (UDG)-mediated cloning strategy."
BASE COUNT
20 a 23 c 23 g 29 t
ORIGIN

Query Match 0.28; Score 37; DB 28; Length 105;
Best Local Similarity 100.0%; Pred. No. 0.034; 0; Indels 0; Gaps 0;
Matches 37; Conservative 0; Mismatches 0;
QY 10241 AATTGCTGACCCAGAGGAGGTTGACGTGAGC 10277
|||||
DB 37 AATTGCTGACCCAGAGGAGGTTGACGTGAGC 1

Search completed: March 29, 2000, 14:49:02
Job time: 41596 sec

PN WO9213102-A.
 CC 06-AUG-1992.
 CC 15-JAN-1992: U00340.
 PR 15-JAN-1991: US-642342.
 PA (GENM-) GENMARK.
 PI Georges M, Massey JM;
 DR WPI: 92-284684/34.
 PT Polymorphic bovine DNA markers - used in genetic identification,
 PS gene mapping, and selective breeding
 PS Table 7: Page 271: 517pp: English.
 CC The sequence is that of a bovine microsatellite sequence obt'd. by
 CC screening a library of bovine MhoI DNA fragments of between
 CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
 CC One out of 50 clones cross-hybridised. Assuming independent
 CC distribution of microsatellites and MhoI sites, the frequency of
 CC (T6)₁₆ >9 microsatellites in the bovine genome is estimated at >100,
 CC 000. The sequence information for ca. 230 such bovine microsatellites
 CC is summarised in the specification and indexed herein (see below).
 CC The sequences upstream and downstream of the microsatellite sequence
 CC were used to generate the required RFLP primers for *in vitro*
 CC amplification of the core sp. microsatellite (using the program
 CC OPTIFIND). The microsatellites may be used to identify individuals,
 CC for parentage testing, and in the genetic mapping of economic trait
 CC loci, or genes involved in the determinism of economically important
 CC traits esp. in cattle, to allow selective breeding.
 CC See also Q03501-34437.
 CC Sequence 44 Bp: 22 A: 0 C: 0 G: 22 T:

Query Match Similarity 0.1%; Score 43; DB 1; Length 44;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY 20222 TTATATATATATATATATATATATATATATATATATA 20264
|||||
DB 44 TTATATATATATATATATATATATATATATATATATA 2
|||||

RESULT	3
ID	Q24810/c
AC	Q24810 standard; DNA; 63 BP.
CD	Q24810;
DE	14-JUL-1992 (first entry)
DT	Oligonucleotide containing Ripe60 binding motif (ATY)15.
DR	Dihydrofolate reductase; stably bent DNA; Ripe60; Ripe100; helicase;
KW	DNA footprinting; (ATY) binding factor; yeast ARS1.
KM	Oligonucleotide affinity column; ss.
OS	Synthetic.
FT	Key
FT	repeat_region
FT	9..53
FT	/*tag- a
FT	9..11
FT	/*tag- b
FT	/note- "ripe60 binding motif"
PN	WO9203479-A.
PD	05-MAR-1992.
PF	27-ANG-1991: U06119.
PR	27-ANG-1990: US-575370.
PA	(UYVE) ROCKEFELLER UNIV.
PI	(UYVE) UNIV OF VERMONT.
PI	Heintz N, Heintz NH, Dailey LA, Caddle MS;
DR	WPI1: 92-068639/12.
PI	Replication initiator protein complex - comprises 60 and 100 kd
PI	iractions for diagnosis, treatment and prevention of sepsis,
PI	viral infection and cancer
PS	Example 2; page 42; 86pp: English.
PC	Completion experiments using the DNaseI protection assay showed
CC	that an oligonucleotide contg. 15 tandem AAT repeats was an
CC	effective competitor for binding of the Hela factor. This oligo was
CC	used to prepare an oligonucleotide affinity column for purification
CC	of Ripe60 from Chinese hamster cells. The protein recognises and
CC	binds to an (ATY)n motif in the gene ori. It can also bind to
CC	the yeast ARS1 domain B. See Q22753 and Q24811.

[illegible]

RESULT 4
071808
ID 071808 standard; DNA; 80 BP.
AC 071808;
AD 30-MAR-1995 (first entry)
DE pSPORT-1a synthetic polyester gene 5' fragment insert.
KW Synthetic gene; synthesis; monomer unit; lactate; glycolate; staples;
KW stop codon; stop signal; monomer unit; lactate; glycolate; staples;
KW translation initiation; modified; Met-tRNA^f; cyanogen bromide; plins;
KW co-polymer; sutures; drug delivery device; screws; ds.
OS Synthetic.
PN 02A112216-1.
PD 01-JUL-1994.
PF 31-DEC-1993; 112716
PR 31-DEC-1992; US-99520.
PA (US50) US SURGICAL CORP.
PI Gruskin EA.
DR WPI; 94-272324/34.
DT New biocompatible medical devices and articles - using a lactic
PT acid-glycolic acid polyester in which monomer units are
PT individually and specifically determined
PS Example 1; Page 45; 67pp; English.
CC This sequence represents the 5' fragment of the synthetic genes which
CC encode polyesters given in 071801-02. This fragment was ligated into
CC the Klenow blunted sequences given in 071808-09 which represent the 3'
CC and 5' ends of the plasmid pSPORT 1b digested with PstI. This places
CC this synthetic fragment under the control of the E. coli trpA initiator.
CC The E. coli trpA initiator sequence was used to express the sequences
CC given in 071801-02 in an E. coli cell free medium. The expressed
CC sequences represent synthetic genes which were used in the synthesis
CC of polyesters derived from the a-amino acid analogues of an amino acid
CC except Pro. Of the three stop codons, UAA, UAG and UGA, one of these
CC stop codons is reversed for the stop signal for the polyester. The
CC remaining two stop codons are available for encoding the monomer
CC units of the polyester. Lactate and glycolate are encoded by UAA and
CC UAG respectively. Initiation of translation always starts at an AUG
CC codon therefore the first lactate of the chain may be encoded by an
CC AUG codon. This is accomplished through the use of a specially modified
CC Met-tRNA^f modified to carry lactate instead of Met. Alternatively, a
CC Met is incorporated in the first position of the polymer chain. The
CC resulting polymer is treated with cyanogen bromide to remove the Met.
CC co-polymers produced in this manner may be used to produce sutures,
CC staples, clips, drug delivery devices, pins and screws.
Sequence 80 BP; 47 A; 6 C; 4 G; 23 T.

```

RESULT      5
Q71814
ID   Q71814 standard; DNA; 81 BP.
AC   Q71814;
DT   30-MAR-1995 (first entry)
DE   Polyester coding sequence for Insertion into pMAL-p2.
Db     1 AATAAATAAATAAATAAATAAATAAATAAATAAATAA 41
QY  27146 AATAAATAAATAAATAAATAAATAAATAAATAAATAA 27186

```

KW Synthetase gene; synthetase; polyester; α -amino acid analogue; clips;
 KW stop codon; stop signal; monomer unit; lactate; glycolate; staples;
 KW translation initiation; modified; Met-tRNA^f; cyanogen bromide; pins;
 KW co-polymer; sutures; drug delivery device; screws; ds.
 OS Synthetic.
 PN CA2112716-A.
 PD 01-JUL-1994.
 PF 31-DEC-1993; 112716.
 PR 31-DEC-1992; US-999520.
 PA (USUS) US SURGICAL CORP.
 PI Grusklin EA;
 PR MPI: 94-272324/34.
 PR New biocompatible medical devices and articles - using a lactic
 PT acid-glycolic acid polyester in which monomer units are
 PT individually and specifically determined
 PS Example 2: Page 53; 67pp; English.
 CC The sequences given in Q71814-16 represent fragments of the
 CC synthetic polyester genes given in Q71801-02. These fragments were
 CC used in the production of the full length polyester gene in the plasmid
 CC pMAL-p2. This plasmid was used so that the polyester could be expressed
 CC in an E. coli cell free medium, as pMAL-p2 is recognised by E. coli
 CC ribozymes and contains the necessary initiator sequences and the
 CC template for the polypeptide leader. The resulting plasmid pMAL-2c
 CC contains the complete polyester coding sequence ligated in frame with
 CC the leader sequence. The expressed sequences represent synthetic
 CC genes which were used in the synthesis of polyesters derived from the
 CC α -amino acid analogues of an amino acid except Pro. Of the three stop
 CC codons, UAA, UAG and UGA, one of these stop codons is reversed for the
 CC stop signal for the polyester. The remaining two stop codons are
 CC available for encoding the monomer units of the polyester. Lactate
 CC and glycolate are encoded by UAA and UAG, respectively. Initiation of
 CC translation always starts at an AUG codon, therefore the first lactate
 CC of the chain may be encoded by an AUG codon. This is accomplished
 CC through the use of a specially modified Met-tRNA^f modified to carry
 CC lactate instead of Met. Alternatively, a Met is incorporated in the
 CC first position of the polymer chain. The resulting polymer is treated
 CC with cyanogen bromide to remove the Met. Co-polymers produced in this
 CC manner may be used to produce sutures, staples, clips, drug delivery
 CC devices, pins and screws.
 SO Sequence 81 BP; 47 A; 6 C; 4 G; 24 T;

[illegible]

01-JUN-1994.
31-DEC-1993; 112716.
31-DEC-1992; US-999520.
(USSU) US SURGICAL CORP.
Gruskin EA;
MPI, 94-272324/34.
New biocompatible medical devices and articles - using a lactic acid-glycolic acid polyester in which monomer units are individually and specifically determined
Example 1: Page 45; 67pp; English.
This sequence represents a fragment of the plasmid pSPORT 1c which contains the 5' fragment of the synthetic genes which encode polyesters, given in Q71801-02, under the control of the E. coli trpA transcription initiator and followed by PstI and HindIII restriction sites. A further fragment of the synthetic genes were ligated into pSPORT 1c to give plasmid pSPORT 1d (see also Q71810-11), with the final fragment being inserted in a third round of ligation to give pSPORT 1e (see also Q71812-13). The E. coli trpA initiator sequence was used to express the full length sequences given in Q71801-02 in an E. coli cell free medium. The expressed sequences represent synthetic genes which were used in the synthesis of polyesters derived from the a-amino acid analogues of an amino acid except Pro. Of the three stop codons, UAA, UAG and UGA, one of these stop codons is reversed for the stop signal for the polyester. The remaining two stop codons are available for encoding the monomer units of the polyester. Lactate and glycolate are encoded by UAA and UAG, respectively. Initiation of translation always starts at an AUG codon, therefore the first lactate or the chain may be encoded by an AUG codon. This is accomplished through the use of a specially modified Met-tRNA^f modified to carry lactate instead of Met. Alternatively, a Met is incorporated in the first position of the polymer chain. The resulting polymer is treated with cyanogen bromide to remove the Met. Co-polymers produced in this manner may be used to produce sutures, staples, clips, drug delivery devices, pins and screws.
Sequence 103 bp; 58 A; 7 C; 9 G; 29 T;

```

Query Match          0.1%; Score 41; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.0035;
Matches    41; Conservative    0; Mismatches    0; Indels    0; Gaps    0

Oy   27146 AATATATATATATATATATATATATATATATATATATATATATATATTA 27186
      |||||||
Db    23  AATAATATATATATATATATATATATATATATATATATATATATATTA 63

RESULT       7
O33621
AC O33621 standard; DNA; 53 BP.
DE 02-FEB-1993 (first entry)
DT Microsatellite sequence from clone GBFSR.
KW PCR; selection; primers; OPRIPIW; breeding; cattle; percentage;
OS genetic mapping; traits; amplification; ss.
PN Bos taurus.
PN W09213102-A.
PD 06-AUG-1992.
PF 15-JAN-1992; U00340.
PR 15-JAN-1991; US-642342.
PA (GEMM-) GENMARK.
PI Georges M, Massey JM;
DR WPI; 92-284684/34.
PT Polymorphic bovine DNA markers - used in genetic identification,
pr gene mapping, and selective breeding
Table 7; Page 178; 51pp; English.
CC The sequence is that of a bovine microsatellite sequence obtd. by
CC screening a library of bovine MbOI DNA fragments of between
CC 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MbOI sites, the frequency of
CC (TG)n >9 microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 250 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence

```


DB 19 GCTTTT...TTTTTTT 57

RESULT 14

AC V05728/c V05728 standard; cDNA; 94 BP.

DE 05-JUN-1998 (first entry)

DE Nucleotide sequence of the 3' portion from clone AJ147_1.

KW Secreted protein; antibody; immunoassay reagent;

KW nutritional supplement; therapeutic activity; murine; calmegln; ds.

OS Homo sapiens.

PN MO9746683-A2.

PD 11-DEC-1997.

PF 06-JUN-1997; U09878.

PR 07-JUN-1996; US-659224.

PA (GENY) GENETICS INST. INC.

PI Bowman M, Evans C, Jacobs K, Lavallie ER, McCly JM,

PI Meisberg D, Racle LA, Spaulding V, Treacy M;

PT Nucleic acids encoding secreted proteins from clones within ATCC

PT 98076 - useful as immuno-modulators, anti-proliferative agents,

PS regulators of cell differentiation and tissue growth, etc

PS Claim 25: Page 67; 99pp; English.

CC The present sequence represents the nucleotide sequence of the 3'

CC portion of clone AJ147_1. The clone was isolated from a human adult

CC testes cDNA library using probe V05755. AJ147_1 is a full length

CC clone encoding a secreted protein. V05727 provides the 5' portion

CC sequence. AJ147_1 shows some homology with murine calmegln, a

CC Ca2+-binding protein that is specifically expressed in spermatogenesis.

CC AJ147_1 may share some of its activity. The nucleic acid can be used for

CC expression of recombinant proteins, as tissue, molecular weight or

CC chromosome markers, indicators of genetic disorders and sources of probes

CC and primers. They can also be used to generate anti-protein or anti-DNA

CC antibodies and as components of interaction trap assays etc. The protein

CC is useful for raising antibodies, as immunoassay reagents and as

CC nutritional supplements. The protein may possibly have any of a great

CC variety of therapeutic activities.

CC Sequence 94 BP; 86 A; 4 C; 4 G; 0 U;

SO

Query Match 0.1%; Score 39; DB 1; Length 94;

Best Local Similarity 100.0%; Pred. No. 0.013;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13599 GCTTTT...TTTTTTT 13637

DB 88 GCTTTT...TTTTTTT 50

RESULT 15

ID V19045

AC V19045 standard; DNA; 40 BP.

DE 28-JUL-1998 (first entry)

DE Alu PCR primer 2.

KW PCR; primer; amplification; Alu repeat sequence; vector;

KW circular yeast artificial chromosome; YAC; ss.

OS Synthetic.

OS Saccharomyces sp.

PN MO9801573-A1.

PD 15-JAN-1998.

PF 09-JUL-1996; U11478.

PR (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Koupina NY, Lariouov VL, Perkins EL, Resnick MA;

PI WPI: 98-110234/10.

PT preparation of yeast artificial chromosomes - by in vivo

PT recombination using vector comprising yeast centromere, marker,

PT yeast telomere and nucleic acid for recombination

PS Example 1; Page 45; 117pp; English.

CC This is the nucleotide sequence for the PCR primer used in the

CC amplification of the Alu repeat sequence, which is used to

CC demonstrate the processes described in the invention. It involves

CC the creation and use of circular yeast artificial chromosome (YAC)

CC to selectively clone specific nucleic acids from a background of

CC mixed nucleic acids by introducing the vector(s) into E. coli cells.

CC They can be used to rapidly isolate human DNA where only a part of the

CC sequence of DNA is known. Using the methods large fragments of DNA can

CC be easily cloned and analysed.

CC Sequence 40 BP; 9 A; 8 C; 19 G; 4 T;

SO

Query Match 0.1%; Score 38; DB 1; Length 40;

Best Local Similarity 100.0%; Pred. No. 0.032;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8613 GAGCTGAGCAGAGAGATCGTTGAACCCGGGAGGCG 8650

DB 2 GAGCTGAGCAGAGAGATCGTTGAACCCGGGAGGCG 39

Search completed: March 30, 2000, 00:58:13

Job time: 76091 sec

Mon Apr 3 08:24:26 2000

us-08-852-495c-1_copy_40000_70000.rng

Page 7



2
1
4

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 17:48:08 ; Search time 340.37 Seconds
(without alignments)
10552.676 Million cell updates/sec

Title: US-08-852-495C-1_COPY_40000_70000
Perfect score: 30001
Sequence: 1 GTGTCACACGAGTTGAGCTA.....AACACAGTTGAGATTTGA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 214294 seqs, 59861574 residues

Word size : 0

Total number of hits satisfying chosen parameters: 349160

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/prodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/prodata/1/ina/6_COMB.seq:*
6: /cgn2_6/prodata/1/ina/PCUTS9_COMB.seq:*
7: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	0.1	39	2	US-08-632-673B-1
2	39	0.1	50	2	US-08-241-465B-11
3	38	0.1	58	3	US-08-417-174-42
4	38	0.1	58	3	US-08-231-65A-40
5	38	0.1	58	4	US-09-007-961-40
6	38	0.1	70	4	US-08-776-944-12
7	38	0.1	75	4	US-08-776-944-13
8	38	0.1	80	2	US-07-920-281C-25
9	37	0.1	37	1	US-08-113-646A-44
10	37	0.1	40	4	US-08-771-624B-1
11	37	0.1	40	4	US-08-440-209-4
12	37	0.1	50	1	US-08-420-443-1
13	37	0.1	55	1	US-08-113-646A-41
14	37	0.1	69	2	US-08-702-344-7
15	37	0.1	69	2	US-08-702-344-22
16	37	0.1	84	2	US-08-664-596B-3
17	37	0.1	84	2	US-08-738-367-3
18	37	0.1	90	2	US-08-677-944-1
19	37	0.1	90	2	US-08-677-944-2
20	37	0.1	98	1	US-08-088-658-42
21	37	0.1	98	4	US-08-471-907A-42
22	37	0.1	105	3	US-08-735-381-2
23	35	0.1	85	1	US-08-130-827-97
24	35	0.1	85	1	US-08-478-675-97
25	34	0.1	84	3	US-08-454-557C-91
26	34	0.1	84	3	US-08-454-557C-91

27	34	0.1	84	4	US-08-340-426D-91	Sequence 91, Appl
28	34	0.1	84	4	US-08-340-426D-91	Sequence 91, Appl
29	34	0.1	84	4	US-08-450-673C-91	Sequence 91, Appl
30	34	0.1	84	4	US-08-450-673C-91	Sequence 91, Appl
31	34	0.1	84	6	PCT-US95-17111A-91	Sequence 91, Appl
32	34	0.1	84	6	PCT-US95-17111A-91	Sequence 91, Appl
33	33	0.1	85	3	US-08-454-557C-92	Sequence 92, Appl
34	33	0.1	85	4	US-08-340-426D-92	Sequence 92, Appl
35	33	0.1	85	4	US-08-450-673C-92	Sequence 92, Appl
36	33	0.1	85	6	PCT-US95-17111A-92	Sequence 92, Appl
37	32	0.1	47	4	US-08-778-494B-114	Sequence 114, Appl
38	32	0.1	50	1	US-08-381-572-20	Sequence 20, Appl
39	32	0.1	50	2	US-08-592-820-20	Sequence 20, Appl
40	32	0.1	51	4	US-08-582-562A-8	Sequence 8, Appl
41	32	0.1	51	4	US-08-778-494B-8	Sequence 8, Appl
42	32	0.1	51	4	US-08-859-998-1373	Sequence 1373, Ap
43	32	0.1	52	4	US-08-778-494B-111	Sequence 111, App
44	32	0.1	55	4	US-08-582-562A-16	Sequence 16, Appl
45	32	0.1	55	4	US-08-778-494B-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-08-632-673B-1/C
Sequence 1, Application US/08632673B
Patent No. 5712098
GENERAL INFORMATION:
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Gnlrke, Andreas
APPLICANT: Thomas, Winston J.
APPLICANT: Dryna, Dennis T.
APPLICANT: Ruddy, David
APPLICANT: Wolff, Roger K.
APPLICANT: Feder, John N.
TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS
TITLE OF INVENTION: DIAGNOSTIC MARKERS AND DIAGNOSTIC METHODS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/632,673B
FILING DATE: 16-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 017957-000410
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-632-673B-1


```

SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/231.565A
FILING DATE: 22-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4124
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 755-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 58
TYPE: NUCLEOTIDE
STRANDEDNESS: DOUBLE
TOPOLOGY: UNKNOWN
MOLECULE TYPE: CDNA
US-08-231-565A-40

```

```
Query Match          0.1% Score 38; DB 3; length 58;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

1 RESULT 5
2 US-09-007-961-40
3 ; Sequence 40: Application US/09007961
4 ; Patent No. 5994523
5 ;
6 ; GENERAL INFORMATION:
7 ; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
8 ; APPLICANT: STEVEN A.
9 ; TITLE OF INVENTION: MELANOMA ANTIGENS AND
10 ; TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
11 ; TITLE OF INVENTION: METHODS
12 ; NUMBER OF SEQUENCES: 43
13 ;
14 ; CORRESPONDENCE ADDRESS:
15 ; ADDRESSEE: MORGAN & FINNEGAN
16 ; STREET: 345 PARK AVENUE
17 ; CITY: NEW YORK
18 ; STATE: NEW YORK
19 ;
20 ; COUNTRY: USA
21 ;
22 ; ZIP: 10154
23 ;
24 ; COMPUTER READABLE FORM:
25 ; MEDIUM TYPE: FLOPPY DISK
26 ; COMPUTER: IBM PC COMPATIBLE
27 ; OPERATING SYSTEM: PC-DOS/MS-DOS
28 ; SOFTWARE: ASCII
29 ;
30 ; CURRENT APPLICATION DATA:
31 ; APPLICATION NUMBER: US/09/007,961
32 ; FILING DATE:
33 ; CLASSIFICATION:
34 ;
35 ; PRIOR APPLICATION DATA:
36 ; APPLICATION NUMBER: 08/231,565
37 ; FILING DATE: 22-APR-1994
38 ; ATTORNEY/AGENT INFORMATION:
39 ; NAME: CAROL M. GRUPPI
40 ; REGISTRATION NUMBER: 37,341
41 ;
42 ; REFERENCE/DOCKET NUMBER: 2026-4124
43 ;
44 ; TELECOMMUNICATION INFORMATION:
45 ; TELEPHONE: (212) 758-4800
46 ; TELEFAX: (212) 751-6849
47 ; TELEX: 421792
48 ;
49 ; INFORMATION FOR SEQ ID NO: 40:
50 ; SEQUENCE CHARACTERISTICS:
51 ; LENGTH: 58

```

```

TYPE: NUCLEOTIDE
STRANDEDNESS: DOUBLE
TOPOLOGY: UNKNOWN
MOLECULE TYPE: CDNA
US-09-007-961-40

```

```

Query Match      0.1%; Score 38; DB 4; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 13600 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 13637
      |||iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii
Db 18 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 55

```

[illegible]


```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/113,646A
: FILING DATE: 31-AUG-1993
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/084,406
: FILING DATE: 10-AUG-1987
: ATTORNEY/AGENT INFORMATION:
: NAME: WILSON, MARY J.
: REGISTRATION NUMBER: 32,955
: REFERENCE/DOCKET NUMBER: 1579-20
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 816-4000
: TELEFAX: (703) 816-4100
: TELEX: 200797 NIXN UR
: INFORMATION FOR SEQ ID NO: 44:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 37 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: RNA (genomic)
: US-08-113-646A-44

```

	Query Match	Score 37;	DB 1; length 37;
	Best Local Similarity	100.0%; Pred. No.	0.0015:
Matches	37;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	13602 TTTTCTTTTTTTTTTTTTTTTTTTTG	13638	
Dd	37 TTTTCTTTTTTTTTTTTTTTTTTTTG	1	

RESULT 10
 US-08-771-624B-1
 : Sequence 1, Application US/08771624B
 : Patent No. 5914230
 : GENERAL INFORMATION:
 : APPLICANT: Liu, Yen Ping
 : APPLICANT: Patel, Rajesh D.
 : APPLICANT: Kurn, Nurith
 : APPLICANT: Lin, Claire
 : APPLICANT: Rose, Samuel J.
 : APPLICANT: Ullman, Edwin F.
 : TITLE OF INVENTION: Homogeneous Amplification and Detection
 : TITLE OF INVENTION: of Nucleic Acids
 : NUMBER OF SEQUENCES: 24
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Behring Diagnostics GmbH, c/o Dade Behring Inc.
 : STREET: 1117 Deerfield Road
 : CITY: Deerfield
 : STATE: Illinois
 : COUNTRY: US
 : ZIP: 60015-0778
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/771.624B
 : FILING DATE: 20-DEC-1996
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 60/009,090
 : FILING DATE: 22-DEC-1995
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Ruzsala, Lois K.
 : REGISTRATION NUMBER: 39,074
 : REFERENCE/DOCKET NUMBER: 1030
 : TELECOMMUNICATION INFORMATION:
 :

TELEPHONE: (847) 267-5366
TELEFAX: (847) 267-6024
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
SPANDENESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomid)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: C-terminal
US-08-771-624B-1

```

Query Match          0.1%; Score 37; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13601 TTTTTCCTTTTTTTTTTTTTTTTTTTTTTTT 13637
      |||||
Db    1 TTTTTCCTTTTTTTTTTTTTTTTTTTTTTTT 37

```

```

US-08-440-209-4
: Sequence 4, Application US/08440209
: Patent No. 592857
: GENERAL INFORMATION:
: APPLICANT: Han, Jang H
: APPLICANT: Spaete, Richard R
: TITLE OF INVENTION: Methods and Compositions for Controlling
: TITLE OF INVENTION: Translation of HCV Proteins
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Wolf, Greenfield, and Sacks P.C.
: STREET: 600 Atlantic Avenue
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02210
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/440,209
: FILING DATE: 12-MAY-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/477,895
: FILING DATE: 06-JUN-1995
: APPLICATION NUMBER: US 08/128,583
: FILING DATE: 28-SEP-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Janluk, Anthony J
: REGISTRATION NUMBER: 29,809
: REFERENCE/DOCKET NUMBER: C0772/7004
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-720-3500
: TELEFAX: 617-720-2441
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 40 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cdna
: HYPOTHETICAL: NO
: ANTI-SENSE: YES
: US-08-440-209-4

```

```

; MEDIUM TYPE: floppy disk
COMPTER: TBM PC compatible

```



```

1      SOFTWARE: PatentIn Release #1.0, Version #1.30
2      CURRENT APPLICATION DATA:
3      APPLICATION NUMBER: US/08/702,344
4      FILING DATE:
5      CLASSIFICATION: 536
6      ATTORNEY/AGENT INFORMATION:
7      NAME: Brown, Scott A.
8      REGISTRATION NUMBER: 32,724
9      TELECOMMUNICATION INFORMATION:
10     TELEPHONE: (617) 498-8224
11     TELEFAX: (617) 876-5851
12     INFORMATION FOR SEQ ID NO: 7:
13     SEQUENCE CHARACTERISTICS:
14     LENGTH: 69 base pairs
15     TYPE: nucleic acid
16     STRANDEDNESS: double
17     TOPOLOGY: linear
18     MOLECULE TYPE: CDNA
19     JS-08-702-344-7

```

```
Query Match Similarity   0.1%; Score 37; DB 2; Length 69;
Best Local Similarity    100.0%; Pred. No. 0.0014;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

MOLECULE TYPE: CDNA
; US-08-702-344-22

```

Query Match      0.1% Score 37: DB 2: Length 69:
Best Local Similarity 100.0% Pred. No. 0.0014:
Matches 37: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 13601 TTTTATTTTATTTTATTTTATTTTATTTTATTTT 13637
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69 TTTTATTTTATTTTATTTTATTTTATTTTATTTT 33

```

Search completed: March 30, 2000, 00:47:59
Job time: 75541 sec

Mon Apr 3 08:24:26 2000

us-08-852-495c-1_copy_40000_70000.rni

Page 8

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 14:49:02 ; Search time 8179.5 Seconds
(without alignments)
13848.486 Million cell updates/sec

Title: US-08-852-495c-1_COPY_40000_70000

Sequence: 1 GGTGACATCAGCTTCAGCTA.....AACACACTTGCAATTGA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : EST:
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_est48:*
83: em_est35:*
84: em_est36:*
85: em_est37:*
86: em_est38:*
87: gb_est49:*
88: gb_est50:*
89: gb_est51:*
90: gb_est52:*
91: gb_est53:*
92: em_est39:*
93: em_est40:*
94: em_est41:*
95: em_est42:*
96: em_est43:*
97: em_est44:*
98: em_est45:*
99: gb_est54:*
100: gb_est55:*
101: em_est46:*
102: gb_est56:*
103: gb_est57:*
104: gb_est58:*
105: gb_est59:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	58	0.2	65	81	B36140 HS-1038-A1-

C	2	53	0.2	91	38	AA780764	ac68f12.s
C	3	51	0.2	92	40	AA916300	on22c60.s
C	4	47	0.2	72	62	A1914923	tr272608.x
C	5	44	0.1	95	24	H61099	yf51c12.r1
C	6	43	0.1	68	28	AA083205	zn26h12.r
C	7	43	0.1	68	28	AA083205	zn26h12.r
C	8	43	0.1	93	44	AA083446	AA0832205
C	9	43	0.1	101	25	N54156	yv63g11.s1
C	10	42	0.1	101	25	N54156	yv63g11.s1
C	11	42	0.1	95	24	H61099	yf51c12.r1
C	12	41	0.1	73	51	AD072757	AD072757
C	13	41	0.1	87	20	D00989	H0MS01971
C	14	41	0.1	98	24	H67549	yv68f10.s1
C	15	41	0.1	102	22	H60869	yv08c10.s1
C	16	41	0.1	102	22	H60869	yv08c10.s1
C	17	40	0.1	72	42	AE659742	yv36h02.x
C	18	40	0.1	82	39	AA844232	h192h08.s
C	19	40	0.1	87	24	AA018576	z6h1h02.s
C	20	40	0.1	97	74	AA018576	z6h1h02.s
C	21	39	0.1	155	100	AQ2812107	AC257107
C	22	39	0.1	58	35	C35772	AC25712
C	23	39	0.1	49	39	A1047660	DXE2P586K
C	24	39	0.1	61	39	AA824000	vq93e09.r
C	25	39	0.1	61	68	HSM002261	AA038300
C	26	39	0.1	71	49	A1638201	ts97e12.x
C	27	39	0.1	51	50	F37855	HSMP07094
C	28	39	0.1	72	40	AA927569	cm71f09.s
C	29	39	0.1	77	64	AA083168	xc15a04.x
C	30	39	0.1	79	44	A1298440	qm90a04.x
C	31	39	0.1	79	44	A1298440	qm90a04.x
C	32	39	0.1	82	50	F34490	HSPD10834
C	33	39	0.1	84	45	A1349814	tsa97a01.x
C	34	39	0.1	84	49	A1627692	h18102.x
C	35	39	0.1	93	33	N29277	yy44b05.r1
C	36	39	0.1	93	33	AA428622	zw75509.s
C	37	39	0.1	96	62	A1915297	tx06f09.x
C	38	39	0.1	96	30	R28974	F0-29D 22 w
C	39	39	0.1	96	30	AA199893	z676h02.s
C	40	39	0.1	100	26	W74144	z603b05.r1
C	41	39	0.1	101	35	AA555145	p107g04.s
C	42	39	0.1	102	45	A1438429	h109s04.x
C	43	39	0.1	102	64	AA027056	g169s11.x
C	44	39	0.1	102	64	AA027056	g169s11.x
C	45	39	0.1	102	64	AA086169	g63014A01
C	46	39	0.1	102	64	AA086169	g63014A01
C	47	39	0.1	102	64	AA086169	g63014A01
C	48	39	0.1	102	64	AA086169	g63014A01
C	49	39	0.1	102	64	AA086169	g63014A01
C	50	39	0.1	102	64	AA086169	g63014A01
C	51	39	0.1	102	64	AA086169	g63014A01
C	52	39	0.1	102	64	AA086169	g63014A01
C	53	39	0.1	102	64	AA086169	g63014A01
C	54	39	0.1	102	64	AA086169	g63014A01
C	55	39	0.1	102	64	AA086169	g63014A01
C	56	39	0.1	102	64	AA086169	g63014A01
C	57	39	0.1	102	64	AA086169	g63014A01
C	58	39	0.1	102	64	AA086169	g63014A01
C	59	39	0.1	102	64	AA086169	g63014A01
C	60	39	0.1	102	64	AA	

ALIGNMENTS

FEATURES
 source
 Fax: (206) 685-7301
 Email: kzachron@u.washington.edu
 Sequence Tagged Connector
 Plate: Ctr 820 row: G column: 11
 Class: BAC ends
 High quality sequence stop: 65.
 Location/Qualifiers
 1..65
 /oranism="Homo sapiens"
 /db_xref="taxon:9606"
 /cclone="Plate-Ctr 820 Col-11 Row-G"
 /clone_lib="CTR Human Genomic Sperm Library C"
 /sex="M"
 /note="Organ: sperm; Vector: pBLOBLAC11; BAC clones in E-Coli DH10b"
 BASE COUNT
 ORIGIN
 25 a 18 c 10 g 12 t
 0.2%; Score 58; DB 81; Length 65;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 5,4e-08;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0
 CY 9842 GGTTAATTTTGTGATTATTTAGTAGAGACGGGGTTTCACCATGTGTGCACGGCTGTCT 9899
 Db 59 GGTTAATTTTGTGATTATTTAGTAGAGACGGGGTTTCACCATGTGTGCACGGCTGTCT 2
 RESULT 2
 AA780764/C 91 bp mRNA EST 05-FEB-1998
 LOCUS ac6ff82.1 Stragene fetal retina j37202 Homo sapiens CDNA clone
 DEFINITION IMAGE:867791 3' similar to contains Alu repetitive element,, mRNA
 sequence.
 accession AA780764
 version AA780764.1 GI:2840095
 keywords EST
 source human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 91)
 Miller,D., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Hillman,D., Kutabch,T., Lacy,K., Le,N., Lennon,G., Maier,M.,

RESULT 1
B36140.AC LOCUS
DEFINITION B36140 65 bp DNA GSS 17-OCT-1997
HS-1038-A1-D06-MR.ab1 C/T Human Genomic Sperm Library C Homo sapiens genomic clone Plate-CT 820 Col-11 Row-G, genomic survey sequence.
ACCESSION B36140
VERSION B36140.1 GI:2535509
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 65)
AUTHORS Mahaliras,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
Tsalicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
TITLE Construction of a Characterized Clone Resource for Genomic Sequencing: Generation and Preliminary Analysis of 20,000 Sequence Tagged Connectors Unpublished (1997)
JOURNAL Contact: Mahaliras Gg, zackrone KD, Hood L
COMMENT University of Washington
Seattle, WA 98195, USA
Tel.: (206) 616-8744

```

FEATURES
    source
        Contact: Wilson RK
        Washington University School of Medicine
        4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
        Tel: 314 286 1800
        Fax: 314 286 1810
        Email: est@watson.wustl.edu
        This clone is available royalty-free through LINL; contact the
        IMAGE Consortium (info@image.llnl.gov) for further information.
        Seq primer: -40m13 fwd. ET from Amer sham.
        Location/Qualifiers
            1..91
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone IMAGE:867791
            /clone_lib="Stratagene fetal retina 937202"
            /sex="mixed"
            /lab_host="50LR (kanamycin resistant)"
            /note="Vector: pBluescript SK+; Site:1: EcoRI; Site:2:
            XhoI; Cloned unidirectionally. Primer: Oligo dr. Pooled
            retinal tissue. Average insert size: 1.0 kb unl. GAP XR
            Vector: 5' adaptor sequence: 5'GCAATTCGCGACGACGAC 3'
            adaptor sequence: 5' CTCGATGTTTCTTTTCTTTT 3'"
BASE COUNT
    15 a      33 c      24 g      19 t
ORIGIN

```

Query Match	0.2%	Score 53:	DB 38:	Length 91:
Best Local Similarity	100.0%	Pred. No. 1.7e-06:		
Matches 53:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
0y 27069	CGGAGGCGGAGGTTGCAGTGCAGCCGAGATTGCGCACCACGTCCAGCTGG	27121		
Db 74	CGGAGGCGGAGGTTGCAGTGCAGCCGAGATTGCGCACCACGTCCAGCTGG	22		
RESULT 3				
LOCUS	AA916300/c			
DEFINITION	on22c08.s1 NCI-CGAP Lu5 Homo sapiens cDNA clone IMAGE:1557422	3'		
ACCESSION	AA916300			
VERSION	AA916300.1	GI:3055692		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
AUTHORS	Eutheria; Primates; Carnivora; Homiidae; Homo.			
TITLE	1 (bases 1 to 92)			
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.			
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),			
	Unpublished (1997)			
	On May 5, 1995 this sequence version replaced gi:797742.			
	Contact: Robert Strausberg, Ph.D.			
	Tel: (301) 496-1550			
	Email: Robert_Strausberg@nih.gov			
	Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.			
	Emmert-Buck, M.D., Ph.D.			
	cDNA Library Preparation: M. Bento Soares, Ph.D.			
	DNA Sequencing by: Washington University Genome Sequencing Center			
	clone distribution: NCI-CGAP clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:			
	www-bio.llnl.gov/db/rp/image/image.html			
FEATURES				
Source	Insert Length: 1519 Std Error: 0.00			
	Seq primer: -40m13 fwd. Err from Amersham			
	High quality sequence stop: 61.			
	Location/Qualifiers			
	1..92			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:1557422"			
	/clone_lib="NCI CGAP Lu5"			
	/tissue_type="carcinoid"			
	/lab_host="DH10B"			
	/note="Organ: Lung; Vector: pT73b-Pac (Pharmacia) with a			
	modified polylinker; 1st strand cDNA was prepared from			
	neuroendocrine lung carcinoid, and was then primed with a			
	Not I - oligo(df) primer. Double-stranded cDNA was ligated			
	to Eco RI adaptors (Pharmacia), digested with Not I and			
	cloned into the Not I and Eco RI sites of the modified			
	pT73 vector. Library is normalized. Library was			
	constructed by Bento Soares and M. Fatima Bonaldo."			
BASE COUNT	15 a 34 c 21 g 22 t			
ORIGIN				
Query Match	0.2%	Score 51:	DB 40:	Length 92:
Best Local Similarity	100.0%	Pred. No. 7e-06:		
Matches 51:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
0y 11995	CAGGAGGCTGAGGAGGAGATTCGTTGAACCCGGAGGTGGAGTTGCAG	12045		
Db 85	CAGGAGGCTGAGGAGGAGATTCGTTGAACCCGGAGGTGGAGTTGCAG	35		

```

RESULT      4          A1914923       72 bp    mRNA           EST           28-JUL-1999
LOCUS       A1914923   tr27e08.x1 NCI-CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2219558 3'
DEFINITION Similar to contains Alu repetitive element; mRNA sequence.
ACCESSION   A1914923
VERSION     A1914923.1 GI:5634778
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Euteria; Primates; Carnivora; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 72)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     On May 1, 1997 this sequence version replaced gi:2059395.
            Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            CGA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLB at:
            www.bio.linnl.gov/bdip/Image/Image.html

FEATURES
Source      Seq primer: -40UP from Gldco.
            Location/Qualifiers
             1..72
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:2219558"
              /clone_1id="NCI-CGAP_OV23"
              /tissue_type="tumor, 5 pooled (see description)"
              /lab_host="DH10B"
              /note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI;
              Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
              Average insert size 1.35 kb. Tumor types include: mixed
              Mullerian tumor, papillary serous, clear cell, spindle
              cell. All are primary tumors, metastasis positive. Life
              Technologies catalog #: 11534-013"
BASE COUNT   8 a                11 c                15 g                38 t
ORIGIN
Query Match      0.2%; Score 47; DB 62; Length 72;
Best Local Similarity 100.0%; Pred. No. 0 00014;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3759 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGCTCCTACTCT 3805
|||||TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
Db 4 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAGCTCCTACTCT 50

RESULT      5          H61099        95 bp    mRNA           EST           06-OCT-1995
LOCUS       H61099/c  yz51c12.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
DEFINITION IMAGE:208822 5', mRNA sequence.
ACCESSION   H61099
VERSION     H61099.1 GI:1013931
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Euteria; Primates; Carnivora; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 95)
AUTHORS     Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,

```

TITLE Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P., and
Wilson, R.
JOURNAL The Masny-Merck EST Project
COMMENT Unpublished (1995)
On Sep 21, 1992 this sequence version replaced g1:279312.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Insert Size: 1755

High quality sequence stops: 87
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1755 Std Error: 0.00
Seq primer: M13Rpl
High quality sequence stop: 87.

FEATURES

Location/Qualifiers

1..95
/organism="Homo sapiens"
/db_xref="GDB:3777953"
/db_xref="taxon:9606"
/clone="IMAGE:208822"
/clone_1lb="Soares fetal liver spleen IMFLS"
/sex="male"
/dev_stage="20 week post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen. Vector: pTR73D (Pharmacia)
with a modified polylinker. Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
15' AACTGGAGATTAATTAAGATCTTTTCTTTTCTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pTR73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT

44 a 4 c 3 g 43 t 1 others

Query Match 0.1%; Score 44; DB 24; Length 95;
Best Local Similarity 100.0%; Pred. No. 0.0011;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 20223 TATG 20266

Db 95 TATG 52

RESULT

6

LOCUS

AA082205

DEFINITION

68 bp mRNA

ACCESSION

AA082205

VERSION

EST

KEYWORDS

EST

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 68)

AUTHORS

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,
Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.

TITLE

Generation and analysis of 280,000 human expressed sequence tags

JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT On Sep 12, 1996 this sequence version replaced g1:1397700.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu

WARNING: There is evidence that suggests that the 384-well parent
plate of this clone contains both human and mouse derived clones.
Thus, the origin of this clone is uncertain. This caution should be
kept in mind should you use this clone.

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 54.

FEATURES

Location/Qualifiers

1..68
/organism="Homo sapiens"
/db_xref="GDB:3927131"
/db_xref="taxon:9606"
/clone="IMAGE:548615"
/clone_1lb="Stratagene neuroepithelium NT2RAMI 937234"
/dev_stage="Ntera-2/RA+MI neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-. Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. NT2
(Ntera-2/cl.d1) precursor cells induced with Retinoid
Acid for 1 week, followed by 3 weeks in mitotic inhibitors
(Replicate #2). Average insert size: 1.1 kb; Uni-ZAP XR
vector. -5' adaptor sequence: 5' GATTCGCGACGAG 3' -3'
adaptor sequence: 5' CTCGAGCTTTTCTTTTCTTTT 3'.

BASE COUNT

36 a 6 c 23 t

Query Match 0.1%; Score 43; DB 28; Length 68;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 20223 TAT 20265

Db 25 TAT 67

RESULT

7

LOCUS

AA082205/c

DEFINITION

68 bp mRNA

ACCESSION

AA082205

VERSION

EST

KEYWORDS

EST

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 68)

AUTHORS

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,
Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.

TITLE

Generation and analysis of 280,000 human expressed sequence tags

JOURNAL

Genome Res. 6 (9), 807-828 (1996)

MEDLINE

97044478

COMMENT

On Sep 12, 1996 this sequence version replaced g1:1397700.
Contact: Wilson RK
Washington University School of Medicine

ORIGIN	
Query Match	0.1%; Score 43; DB 25; Length 101;
Best Local Similarity	100.0%; Pred. No. 0.0021;
Matches	43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

RESULT 10					
NS4156/c	NS4156	101 bp	mRNA	EST	28-JAN-1997
LOCUS	YV65911.s1	Soares fetal liver	spleen	INRFS	Homo sapiens cDNA clone
DEFINITION	IMAGE:247460	3' similar to contains element	MER1	repetitive element	
	..	mRNA sequence.			

ACCESSION	N54156	
VERSION	N54156.1	GI:1195322
KEYWORDS	EST.	
SOURCE	human.	

ORGANISM

REFERENCE AUTHORS

TITLE
Generation and analysis of 280,000 human expressed sequence tags
JOURNAL
Genome Res. 6 (9), 807-828 (1996)
MEDLINE
97044478
COMMENT
On Nov 22, 1995 this sequence version replaced g1:1070798.

Email: est@watson.wustl.edu
This clone is available royalty-free through LML; contact the
IMAE Consortium (info@lml.gov) for further information.
Insert length: 636 Std Error: 0.00
Seq primer: ml3 -40 forward
High quality sequence stop: 89.

FEATURES

SOURCE

```

/organism="Homo sapiens"
/db_xref="GDB:3796706"
/db_xref="taxon:9606"
/clone="IMAGE:247460"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pRT3D (pharmacia)
with a modified polylinker. Site.1: Pac I; Site.2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer[5'
AGTCGAGAGAATTATTTAAGACACTTTTTTTTTTTTTTTTCTT3']
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pRiv vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

```

Query Match	0.18;	Score 43;	DB 25;	Length 101;
Best Local Similarity	100.08;	Pred. No. 0.0021;		

Matches	43;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

[illegible]

RESULT	11
H61099	
LOCUS	H61099 95 bp mRNA EST 06-OCT-1995
DEFINITION	Yr5c1c2_r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE	J208822 5', mRNA sequence.
H61099	
ACCESSION	

```

VERSION      H61099.1  GI:1013931

```

SOURCE
ORGANISATION

REFERENCE

AUTHORS

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1860
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1755
High quality sequence strops: 87
Source: IMAG Consortium, LINT.
This clone is available royalty-free through LINT; contact the
IMAG Consortium (info@image.lint.gov) for further information
Insert Length: 1755 Std Error: 0.00
Seq primer: M13p1

FEATURES
SOURCE

```

/organism="Homo sapiens"
/db_xref="GDB:3777953"
/db_xref="taxon:9606"
/clone="IMAGE:208822"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week.post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: Liver and Spleen; Vector: pTR3D (Pharmacia) with a modified polylinker; Site1: Pac I; Site.2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5'-ACACGAGACAAATTAATTAATGACGCTCTTTTCTTTTCTTTT 3'] double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pTR73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fátima Bonaldo."

```

Query Match	0.18;	Score 42;	DB 24;	Length 95;
Best Local Similarity	100.08;	Pred. No. 0.0046;		
Matches 42;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]


```

RESULT   12
LOCUS     AU072757
DEFINITION AU072757 Dictyostellium discoidem SS (H.Urushihara) Dictyostellium
ACCESSION AU072757
VERSION    AU072757
KEYWORDS   EST.
SOURCE      Urushihara,H.
ORGANISM   Dictyostelium discoidem.
REFERENCE  1 (bases 1 to 73)
AUTHORS    Urushihara,H.
TITLE       Developmental cDNA In Dictyostellum discoidem (1999)
JOURNAL    Unpublished (1999)
COMMENT     On Jun 5, 1998 this sequence version replaced gi:3188495.
            Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            3-3-10 Ten-noda,Tsukuba,Ibaraki 305, Japan
            Email: d402nhesakura.cc.tsukuba.ac.jp
PROJECT - Dictyostelium discoidem CDNA project in Japan.

FEATURES             source
source               /organism="Dictyostelium discoidem"
                   /strain="AX4"
                   /db_xref="taxon:44689"
                   /clone="SSN765"
                   /clone_1lb="Dictyostelium discoidem SS (H.Urushihara)"
                   /dev_stage="slug"
BASE COUNT          46 a                1 g        26 t
ORIGIN              0 c

Query Match         0.1%; Score 41; DB 51; Length 73;
Best Local Similarity 100.0%; Pred.No. 0.011;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 27146 AATATATTAATAAATAATATATATATATATATATATATAA 27186
Db 15 AATATTAATAATTAATTAATTAATTAATTAATTAATTAATTA 55

RESULT   13
LOCUS     D20989/c
DEFINITION HONGSO1971 Human promyelocyte Homo sapiens CDNA clone mp0383 3',
ACCESSION D20989
VERSION    D20989
KEYWORDS   EST.
SOURCE      human.
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 87)
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homilidae; Homo.
TITLE       Yoshinari,H., Arimoto,J. and Matsubara,K.,
            Yoshinari,H., Arimoto,J. and Matsubara,K.,
            Gene expression of human promyelocytic cell line HL60 before and
            after induction of differentiation. A new application of 3'directed
            cDNA sequencing
            unpublished (1993)
JOURNAL    Contact: Okubo,K., Fukushima,A., Yoshii,J., Niiyama,T., Kojima,Y.,
COMMENTS   Yoshinari,H., Arimoto,J. and Matsubara,K.,
            Institute for Molecular and Cellular Biology
            Osaka University
            3-1 Yamadaoka,Suita,Osaka 565,Japan.
            Location/Qualifiers
            1..87
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="mp0383"
FEATURES             source
source

```

```

/clone_11b-"Human promyelocyte
/clone-"Female, adult, cell_line = HL60, cell_type =
promyelocyte."
Query Match 0.1%; Score 41; DB 20; Length 87;
Best Local Similarity 100.0%; Prod. No. 0.0099;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20914 GTGGCTCACACCTGTATCCAGCAGCTTTGGAGGCTGAGG 20954
|||||
DB 51 GTGGCTCACACCTGTATCCAGCAGCTTTGGAGGCTGAGG 11
|||||
RESULT 14
LOCUS H67549 98 bp mRNA EST 11-DEC-1995
DEFINITION YU86f10.s1 Weizmann Olfactory Epithelium Homo sapiens cDNA clone
IMAGE:238987.3 similar to contains Alu repetitive element;contains
MEN22 repetitive element ;, mRNA sequence.
ACCESSION H67549
VERSION H67549.1 GI:1114431
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 98)
AUTHORS Hillier,L., Lennon,G., Becker,M., Donald,M.F., Chiapelli,B.,
Chissee,S., Dietrich,N., Dubucque,T., Favellio,A., Gish,W.,
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
Roehling,T., Scheilenderg,K., Soares,M.B., Tan,F., Tillery-Meg,J.,
Trevaskis,E., Underwood,R., Wohlmann,P., Waterston,R., Wilson,R.
and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
COMMENT
On Oct 18, 1995 this sequence version replaced gi:1026289.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estevenson.wustl.edu
Insert Size: 174
Source: IMAGE Consortium, LNLN
This clone is available royalty-free through LNLN ; contact the
IMAGE Consortium (info@image.lnl.n.gov) for further information.
Seq primer: Promega -21ml3.
Location/Qualifiers
1..98
/organism="Homo sapiens"
/db_xref="GDB:3864371"
/db_xref="taxon:9606"
/clone="IMAGE:238987"
/clone_lib="Weizmann Olfactory Epithelium"
/sex="female"
/tissue_type="olfactory epithelium"
/dev_stage="35 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: nose; Vector: pBluescript SK-; Site_1:
EcoRI, Site_2: XhoI; Cloned unidirectionally. Primer:
0.150 dt. Olfactory epithelium, normal. Average insert
size: 0.8 kb; Uni-Zap XR vector. Library constructed by N.
Walker, D. Lancelot, Weizmann Institute of Science. -5'
adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTGTTTTGTTTT 3'."
BASE COUNT 19 a 39 c 23 g 17 t
ORIGIN

```

Mon Apr 3 08:24:27 2000

us-08-852-495c-1_copy_40000_70000.rst

Page 8

Query Match 0.1%; Score 41; DB 24; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.0092;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9768 CCTCCAGGTTCAGCAATTCCTGCTCAGCCGCCAG 9808
|||||
DB 50 CCTCCAGGTTCAGCAATTCCTGCTCAGCCGCCAG 90

Db 50 ATATATATATATATATATATATATATATATATATATATA 90
Search completed: March 29, 2000, 20:38:58
Job time: 62592 sec

RESULT 15

R60869

LOCUS

DEFINITION

IMAGE:42595 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Apr 18, 1995 this sequence version replaced gi:775532.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

Insert Size: 1405

High quality sequence stops: 76 Source: IMAGE Consortium, LNL. This

clone is available royalty-free through LNL; contact the IMAGE

Consortium (info@image.llnl.gov) for further information.

Insert Length: 1405 Std Error: 0.00

Seq primer: SP6

High quality sequence stop: 76.

Location/Qualifiers

1..102

/organism="Homo sapiens"

/db_xref="GDB:415186"

/db_xref="taxon:9606"

/clone_image:42595"

/clone_lib="Soares Infant brain INIB"

/sex="female"

/dev_stage="73 days post natal"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: whole brain; Vector: lafmid BA; Site:1: Not

I; Site:2: Hind III; 1st strand cDNA was primed with a Not

I - oligo(dT) primer [5',

AACTGAGAGATTCGCGCCGCGAGCAATTTTCTTTTCTTTT 3'];

double-stranded cDNA was ligated to Hind III adaptors

(Pharmacia), digested with Not I and directionally cloned

into the Not I and Hind III sites of the lafmid BA vector.

Library went through one round of normalization. Library

constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 35 a 9 c 5 g 52 t 1 others

ORIGIN

Query Match 0.1%; Score 41; DB 22; Length 102;

Best Local Similarity 100.0%; Pred. No. 0.0091;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20224 ATATATATATATATATATATATATATATATATATA 20264

|||||

Mon Apr 3 08:24:27 2000

us-08-852-495c-1_copy_40000_70000.rst

Page 9



•

•

•

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 00:22:34 ; Search time 13413.1 Seconds

(Without alignments)
-6791.482 Million cell updates/sec

Title: US-08-852-495c-1_COPY_65000_95000

Perfect score: 30001
Sequence: 1 TTATATATGACAAATTAAT.....TACAGATCTCTGCAGTTT 30001Scoring table: OLIGO_NUC
Gapop 60.0 , Gapept 60.0

Searched: 821193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pl4:*
11: gb_pl5:*
12: gb_pl6:*
13: gb_pl7:*
14: gb_pl8:*
15: gb_pl9:*
16: gb_pl10:*
17: gb_pl11:*
18: gb_pl12:*
19: gb_pl13:*
20: gb_pl14:*
21: gb_pl15:*
22: gb_pl16:*
23: gb_pl17:*
24: gb_pl18:*
25: gb_pl19:*
26: gb_pl20:*
27: gb_pl21:*
28: gb_pl22:*
29: gb_pl23:*
30: gb_pl24:*
31: gb_pl25:*
32: gb_pl26:*
33: gb_pl27:*
34: gb_pl28:*
35: gb_pl29:*
36: gb_pl30:*
37: gb_pl31:*
38: gb_pl32:*
39: gb_pl33:*
40: gb_pl34:*
41: gb_pl35:*
42: gb_pl36:*
43: gb_pl37:*

44: gb_hlg6:*
45: gb_hlg7:*
46: em_hlg1:*
47: em_hlg2:*
48: em_hlg3:*
49: em_hlg4:*
50: gb_pl13:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	0.2	76	5 AR051499	AR051499 Sequence
2	55	0.2	60	5 AR051487	AR051487 Sequence
3	48	0.2	48	9 HUMH116	D28457 Human mRNA
4	41	0.1	104	9 HUMALICE272	M87889 Human carcl
5	36	0.1	57	13 HUMUT1543A	L30907 Human STS U
6	36	0.1	66	40 AF087511	AF087511 Homo sapi
7	35	0.1	57	13 HUMUT1543A	L30907 Human STS U
8	35	0.1	85	5 AR051522	AR051522 Sequence
9	34	0.1	42	5 HUMALUANC	L36838 Homo sapien
10	34	0.1	84	5 AR051521	AR051521 Sequence
11	33	0.1	35	5 A25212	A25212 Inter-Alu s
12	33	0.1	35	24 E09140	E09140 Synthetic D
13	33	0.1	40	5 A68621	A68621 Sequence 1
14	33	0.1	65	9 HUMALUANC	L36836 Homo sapien
15	33	0.1	69	13 HUMUT578A	L39139 Human STS U
16	33	0.1	85	5 AR051522	AR051522 Sequence
17	32	0.1	51	10 S62605	S62605 Cl-Inhibito
18	32	0.1	63	3 DOGSR7SL1	M30845 Dog signal
19	32	0.1	67	9 HUMALUANC	L36843 Homo sapien
20	32	0.1	80	9 HUMBRKFC	M36133 Human alpha
21	31	0.1	66	10 S54071	S54071 hypoxanthin
22	31	0.1	67	13 HUMUT44A	L29843 Human STS U
23	31	0.1	90	9 HUMDLRPL	K03535 Human low d
24	31	0.1	90	9 HUMDLRPL	K03535 Human low d
25	31	0.1	90	9 HUMDLRPL	M15365 Human low d
26	31	0.1	100	11 HSU67848	U67848 Human beta-
27	31	0.1	101	13 HUMUT832A	L31225 Human STS U
28	31	0.1	104	5 A08899	A08899 H.sapiens (
29	31	0.1	40	5 A68621	A68621 Sequence 1
30	30	0.1	66	40 AF087511	AF087511 Homo sapi
31	30	0.1	67	13 HUMUT44A	L29843 Human STS U
32	30	0.1	73	5 A08915	A08915 H.sapiens (
33	30	0.1	80	9 HUMBRKFC	M36134 Human alpha
34	30	0.1	80	9 HUMBRKFC	M36135 Human alpha
35	30	0.1	91	13 G38859	G38859 TA50 Plasm
36	30	0.1	95	4 GGRE10H7	X78616 G.gallus ge
37	30	0.1	99	13 HUMUT7692A	L30306 Human STS U
38	30	0.1	101	13 HUMUT832A	L31225 Human STS U
39	30	0.1	102	5 A08911	A08911 H.sapiens (
40	30	0.1	103	9 HUMALICE221	M87886 Human carcl
41	30	0.1	105	13 G32655	G32655 A009L30 Hum
42	30	0.1	105	13 HS015XB3	Z50878 H.sapiens (
43	29	0.1	60	5 AR051490	AR051490 Sequence
44	29	0.1	63	13 HUMUT5302A	L30829 Human STS U
45	29	0.1	65	9 HUMALUANC	L36836 Homo sapien

ALIGNMENTS

RESULT	1				
LOCUS	AR051499		76 bp	DNA	
DEFINITION	Sequence	69	from patent US 5830670.		
ACCESSION	AR051499				
VERSION	AR051499.1		GI:5974863		
				PAT	29-SEP-1999

KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 76)
AUTHORS de la Monte,S., and Wands,J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 69 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..76
/organism="unknown"
BASE COUNT 15 a 19 c 18 g 24 t
ORIGIN

Query Match 0.2%; Score 58; DB 5; Length 76;
Best Local Similarity 100.0%; Pred. No. 7.8e-19;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27685 GCCCGCGTAATTTTGTATTTAGTAGACAGAGGTTTCACCGTTGGCCAGATG 27742
DB 6 GCCCGCGTAATTTTGTATTTAGTAGACAGAGGTTTCACCGTTGGCCAGATG 63

RESULT 2
LOCUS AR051487 60 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 57 from patent US 5830670.
ACCESSION AR051487
VERSION AR051487.1 GI:5974851
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 60)
AUTHORS de la Monte,S., and Wands,J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 57 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..60
/organism="unknown"
BASE COUNT 12 a 14 c 15 g 19 t
ORIGIN

Query Match 0.2%; Score 55; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.9e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27685 GCCCGCGTAATTTTGTATTTAGTAGACAGAGGTTTCACCGTTGGCCAGG 27739
DB 6 GCCCGCGTAATTTTGTATTTAGTAGACAGAGGTTTCACCGTTGGCCAGG 60

RESULT 3
LOCUS HUMH116 48 bp mRNA PRI 07-FEB-1999
DEFINITION Human mRNA for histone H1D, 5'UTR (sequence from the 5'cap to the start codon).
ACCESSION D28457
VERSION D28457.1 GI:461197
AUTHORS histone H1D.
KEYWORDS Homo sapiens (library:HT-1080/pkA1) fibrosarcoma cell-line HT-1080
SOURCE cDNA to mRNA, clone HP00474.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.
REFERENCE Kato,S.
AUTHORS Kato,S.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-1994) to the DDBJ/EMBL/GenBank databases. Setshl

Kato, Sagami Chemical Research Center, Genetic Engineering Section;
4-4-1 Nishi-Onuma, Sagamihara, Kanagawa 229, Japan
(E-mail:btm00121@biotech.net.com, Tel:0427-42-4791(ex.415),
Fax:0427-49-7631)
REFERENCE 2 (bases 1 to 48)
AUTHORS Kato,S.
JOURNAL Unpublished (1994)
REFERENCE 3 (sites)
AUTHORS Kato,S., Sekine,S., Oh,S.W., Kim,N.S., Umezawa,Y., Abe,N.,
Yokoyama-Kobayashi,M. and Aoki,T.
TITLE Construction of a human full-length cDNA bank
JOURNAL Gene 150 (2), 243-250 (1994)
MEDLINE 95121910
COMMENT Submitted (03-Feb-1994) to DDBJ by:
Seishi Kato
Genetic Engineering Section
Sagami Chemical Research Center
4-4-1 Nishi-Onuma
Sagamihara, Kanagawa 229
Japan
Phone: 0427-42-4791
Fax: 0427-49-7631.

FEATURES
source 1..48
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HT-1080"
/clone_id="HT-1080/pkA1"
/tissue_type="fibrosarcoma"
1..42
43..>48
/product="histone H1D"
/protein_id="BA05823.1"
/db_xref="GI:433162"
/translation="MS"
BASE COUNT 8 a 14 c 9 g 17 t
ORIGIN

Query Match 0.2%; Score 48; DB 9; Length 48;
Best Local Similarity 100.0%; Pred. No. 1.4e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12760 ATCGGCGTTGGCAGCTTGACACCGAGTTTGTGATCTCAACATGTCC 12807
DB 1 ATCGGCGTTGGCAGCTTGACACCGAGTTTGTGATCTCAACATGTCC 48

RESULT 4
LOCUS HUMALCE272/C 104 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE272.
ACCESSION M87889
VERSION M87889.1 GI:174875
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 104)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES Location/Qualifiers
source 1..104
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="M8782D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"

```
BASE COUNT      22 a      26 c      37 g      19 t
ORIGIN
Query Match      0.1%; Score 41; DB 9; Length 104;
Best Local Similarity 100.0%; Pred. No. 7.3e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2681 CCTCAGCCTCCCAAGTGTGGATTACAGAGTGAGCCAC 2721
      |||||||
Db 56 CCTCAGCCTCCCAAGTGTGGATTACAGAGTGAGCCAC 16

RESULT 5
HUMUT1543A/c 57 bp DNA SRS 17-JAN-1995
LOCUS HUMUT1543A
DEFINITION Human SRS UT1543, 5' primer bind, sequence tagged site.
ACCESSION L30907
VERSION L30907.1 GI:624443
KEYWORDS SRS; PCR primer; SRS sequence; microsatellite DNA; microsatellite
SOURCE marker; sequence tagged site; tetranucleotide repeat.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS 1 (bases 1 to 57)
Gerken,S.C., Matsunami,N., Plaetke,R., Albertsen,H., Ballard,L.,
Melis,R., Lawrence,E., Moore,M., Holik,P.R., Carlson,M., Zhao,X.,
Robertson,M., Bradley,P., Elsner,T., Tingey,A., Lalouel,J.-M. and
White,R.
TITLE Genetic and physical mapping of simple sequence repeat containing
JOURNAL sequence tagged sites from the human genome
COMMENT Unpublished (1994)
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: steecorona.med.utah.edu
Primer A: CGAGACTCCGTCACAGAAA
Primer B: CCATCCACTCTTCATG
End to Label: Primer A
PCR Profile:
Initial Denaturation: 94C 300sec
Cycles Denaturation Annealing Extension 5 94
C 10 sec. 56 C 10 sec. 72 C 20 sec. 30 94 C 10 sec.
56 C 10 sec. 72 C 20 sec. Mg++: 1.00 mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 5.
FEATURES
source 1..57
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
32..50
primer_bind /evidence=experimental 7 t
BASE COUNT 21 a 16 c 13 g
ORIGIN

Query Match 0.1%; Score 36; DB 13; Length 57;
Best Local Similarity 100.0%; Pred. No. 3.1e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8625 GAGCTCGCTGTGACCCAGCGTGAGTGCAGTGG 8660
      |||||||
Db 39 GAGCTCGCTGTGACCCAGCGTGAGTGCAGTGG 4

RESULT 6
AF087511 66 bp mRNA PRI 13-SEP-1999
LOCUS AF087511/c
DEFINITION Homo sapiens clone ENAc+22 epithelial sodium channel alpha subunit
(SCNN1A) mRNA, alternatively spliced, partial sequence.
ACCESSION AF087511

VERSION AF087511.1 GI:5870626
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
REFERENCE Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS 1 (bases 1 to 66)
Oh,Y. and Warnock,D.G.
TITLE An Alu cassette in the human epithelial sodium channel
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 66)
Oh,Y. and Warnock,D.G.
AUTHORS Direct Submission
TITLE Submitted (26-AUG-1998) Medicine, UAB, 1720 7th Ave. SO.,
JOURNAL Birmingham, AL 35294, USA
FEATURES
source 1..66
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="ENAc+22"
<1..>66
/gene="SCNN1A"
/note="alternatively spliced epithelial sodium channel
alpha subunit containing intronic Alu sequence"
<1..>66
/gene="SCNN1A"
BASE COUNT 12 a 20 c 24 g 10 t
ORIGIN

Query Match 0.1%; Score 36; DB 40; Length 66;
Best Local Similarity 100.0%; Pred. No. 3.1e-07;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28853 CCGCCCTCGGCTCCCAAGTGTGGATTACAGT 28888
      |||||||
Db 55 CCGCCCTCGGCTCCCAAGTGTGGATTACAGT 20

RESULT 7
HUMUT1543A 57 bp DNA SRS 17-JAN-1995
LOCUS HUMUT1543A
DEFINITION Human SRS UT1543, 5' primer bind, sequence tagged site.
ACCESSION L30907
VERSION L30907.1 GI:624443
KEYWORDS SRS; PCR primer; SRS sequence; microsatellite DNA; microsatellite
SOURCE marker; sequence tagged site; tetranucleotide repeat.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS 1 (bases 1 to 57)
Gerken,S.C., Matsunami,N., Plaetke,R., Albertsen,H., Ballard,L.,
Melis,R., Lawrence,E., Moore,M., Holik,P.R., Carlson,M., Zhao,X.,
Robertson,M., Bradley,P., Elsner,T., Tingey,A., Lalouel,J.-M. and
White,R.
TITLE Genetic and physical mapping of simple sequence repeat containing
JOURNAL sequence tagged sites from the human genome
COMMENT Unpublished (1994)
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: steecorona.med.utah.edu
Primer A: CGAGACTCCGTCACAGAAA
Primer B: CCATCCACTCTTCATG
End to Label: Primer A
PCR Profile:
Initial Denaturation: 94C 300sec
Cycles Denaturation Annealing Extension 5 94
C 10 sec. 56 C 10 sec. 72 C 20 sec. 30 94 C 10 sec.
56 C 10 sec. 72 C 20 sec. Mg++: 1.00 mM
```

Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 5.
FEATURES
Source Location/Qualifiers
1..57 /organism="Homo sapiens"
/db_xref="taxon:9606"
primer_bind 32..50
/evidence=experimental
BASE COUNT 21 a 16 c 13 g 7 t
ORIGIN

Query Match 0.1%; Score 35; DB 13; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26376 GCACCTCCAGCCTGGGTGACAGCGAGACTCCGTC 26410
Db 9 GCACCTCCAGCCTGGGTGACAGCGAGACTCCGTC 43

RESULT 8
LOCUS AR051522 85 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 92 from patent US 5830670.
ACCESSION AR051522
VERSION AR051522.1 GI:5974886
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 85)
AUTHORS de la Monte, S. and Wands, J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 92 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..85
/organism="unknown"
BASE COUNT 13 a 27 c 25 g 20 t
ORIGIN

Query Match 0.1%; Score 35; DB 5; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21802 CGGGGGATCAGAGGTGACGATCGACACATC 21836
Db 48 CGGGGGATCAGAGGTGACGATCGACACATC 14

RESULT 9
LOCUS HUMALVANC 42 bp DNA PRI 08-OCT-1994
DEFINITION Homo sapiens 4000 year old remains from Nekht-anhk Alu repeat fragment 12:4.
ACCESSION L36838
VERSION L36838.1 GI:556195
KEYWORDS Alu repeat.
SOURCE Homo sapiens (individual isolate 4000 year old remains from Nekht-anhk) liver DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Euteria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (sites)
AUTHORS Paabo, S.
TITLE Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (6), 1939-1943 (1989)
MEDLINE 89184542
FEATURES Location/Qualifiers
source 1..42

/organism="Homo sapiens"
/isolate="4000 year old remains from Nekht-anhk"
/db_xref="taxon:9606"
/tissue_type="liver"
repeat_region <1..>42
/rpt_family="Alu"
BASE COUNT 8 a 10 c 17 g 7 t
ORIGIN

Query Match 0.1%; Score 34; DB 9; Length 42;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 28846 TGATCCACCGCTCGGCTCCCAAGTGTGG 28879
Db 34 TGATCCACCGCTCGGCTCCCAAGTGTGG 1

RESULT 10
LOCUS AR051521 84 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 91 from patent US 5830670.
ACCESSION AR051521
VERSION AR051521.1 GI:5974885
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 84)
AUTHORS de la Monte, S. and Wands, J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 91 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..84
/organism="unknown"
BASE COUNT 15 a 29 c 21 g 19 t
ORIGIN

Query Match 0.1%; Score 34; DB 5; Length 84;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2676 GCCCGCTCAGCTCCCAAGTGTGGATTACA 2709
Db 44 GCCCGCTCAGCTCCCAAGTGTGGATTACA 77

RESULT 11
LOCUS A25212 35 bp DNA PAT 11-APR-1995
DEFINITION Inter-Alu specific primer DNA (pdj33) from patent WO9213101.
ACCESSION A25212
VERSION A25212.1 GI:904592
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 35)
AUTHORS
TITLE METHOD OF DETECTING DNA SEQUENCE VARIATION
JOURNAL Patent: WO 9213101-A 3 06-AUG-1992;
FEATURES Location/Qualifiers
source 1..35
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 8 a 10 c 11 g 6 t
ORIGIN

Query Match 0.1%; Score 33; DB 5; Length 35;

Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1886 TGGCTCAGCGCTGTAATCCAGCACTTTGGAG 1918
|||||
Db 35 TGGCTCAGCGCTGTAATCCAGCACTTTGGAG 3

RESULT 12

E09140/C
ID E09140 standard: DNA; UNC; 35 BP.

AC E09140;

XX E09140.1

DT 07-OCT-1997 (Rel. 52, Created)

DT 08-OCT-1997 (Rel. 52, Last updated, Version 1)

DE Synthetic DNA for Alu specific primer.

XX JP 1995115999-A/3.

OS unidentified

OC unclassified.

XX [1]

RP 1-35

RA Andoreasu H.A., Yan F.;

RT "DETECTING METHOD FOR DNA ARRANGEMENT VARIATION";

RL Patent number JP 1995115999-A/3, 09-MAY-1995.

RL INGENIT BV.

XX OS None

CC OC Artificial sequences.

CC PN JP 1995115999-A/3

CC PD 09-MAY-1995

CC PF 22-MAY-1992 JP 1992130668

CC PI ANDOREASU HERARUDONUS ARTSUTERURINDEN, YAN FUEIKU

CC PC C1201/68, C12N15/00, G01N27/447, G01N27/447;

CC CC strandedness: Single;

CC CC topology: Linear;

CC CC Key Location/Qualifiers

CC FT source 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

CC FT misc-feature 1..35

KEYWORDS

SOURCE unidentified.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 40)

AUTHORS Resnick M.A., Larionov V.L., Koupina N.Y. and Perkins E.L.

TITLE TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING

JOURNAL Patent: WO 9801573-A 15-JAN-1998;

US HEALTH (US)

FEATURES Location/Qualifiers

source 1..40

/db_xref="taxon:32644"

BASE COUNT 7 a 12 c 13 g 8 t

ORIGIN

Query Match 0.1%; Score 33; DB 5; Length 40;

Best Local Similarity 100.0%; Pred. No. 1.2e-05;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1886 TGGCTCAGCGCTGTAATCCAGCACTTTGGAG 1918

|||||
Db 3 TGGCTCAGCGCTGTAATCCAGCACTTTGGAG 35

RESULT 14

HUMALUANC

LOCUS HUMALUANC

DEFINITION Homo sapiens 4000 year old remains from Nekht-anhk Alu repeat 12.2.

ACCESION L36836

VERSION L36836.1 GI:556194

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

FEATURES

source

repeat-region

BASE COUNT

ORIGIN

Query Match 0.1%; Score 33; DB 9; Length 65;

Best Local Similarity 100.0%; Pred. No. 1.2e-05;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26168 TTTGGAGCGCGAGCGGCGATCAGAGTC 26200

|||||
Db 33 TTTGGAGCGCGAGCGGCGATCAGAGTC 65

RESULT 15

HUMUT578A/C

LOCUS HUMUT578A/C

DEFINITION Human SFS UT578, 5' primer bind, sequence tagged site.

ACCESION L39139 L18336

VERSION L39139.1 GI:642101

KEYWORDS

SFS; PCR primer; SFS sequence; microsatellite DNA; microsatellite

marker; sequence tagged site; tetranucleotide repeat.

Mon Apr 3 08:24:28 2000

us-08-852-495c-1_copy_65000_95000.rge

Page 6

SOURCE Homo sapiens DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 69)
AUTHORS Gerken,S.C., Matsunami,N., Plaetke,R., Albertsen,H., Ballard,L.,
Wells,R., Lawrence,E., Moore,M., Holik,P.R., Carlson,W., Zhao,X.,
Robertson,M., Bradley,P., Eisner,T., Tingey,A., Lalouel,J.-M. and
White,R.
TITLE Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
JOURNAL Unpublished (1994)
COMMENT Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: stecorona.med.utah.edu
Primer A: AGTGCAGACGCGCTGGC
Primer B: AGTGGCAGAAAATCGCATC
End to Label: Primer A
PCR Profile:
Initial Denaturation: 94C 300sec
Cycles Denaturation Annealing Extension 5 94
C 10 sec. 62 C 10 sec. 72 C 20 sec. 30
58 C 10 sec. 72 C 20 sec. Mg++: 1.50 mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 1
FEATURES
source location/Qualifiers
1..69
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19"
primer_bind 14..32
/note="STS UT578 5' end"
/evidence=experimental
BASE COUNT 26 a 16 c 14 g 13 t
ORIGIN
Query Match 0.1%; Score 33; DB 13; Length 69;
Best local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15491 GTTGGCCAGGCTGCTGCAACTCTGACCTCA 15523
|||
DB 36 GTTGGCCAGGCTGCTGCAACTCTGACCTCA 4

Search completed: March 30, 2000, 07:33:31
Job time: 101360 sec

[illegible]

	CC	rbozymes and contains the necessary initiator sequences and the template for the polypeptide leader.
	CC	The resulting plasmid pMAL-2c contains the complete polyester coding sequence ligated in frame with the leader sequence.
	CC	The expressed sequences represent synthetic genes which were used in the synthesis of polyesters derived from the amino acid analogues of an amino acid except pro.
	CC	Of the three stop codons, UAA, UAG and UGA, one of these stop codons is reversed for the stop signal for the polyester.
	CC	The remaining two stop codons are available for encoding the monomer units of the polyester.
	CC	Lactate and glycolate are encoded by UUA and UUG, respectively.
	CC	Initiation of translation always starts at an AUG codon,
	CC	therefore the first lactate or the chain may be encoded by an AUC codon.
	CC	This is accomplished through the use of a specially modified Met-tRNA ^{fMet} modified to carry lactate instead of met.
	CC	Alternatively a Met is incorporated in the first position of the polymer chain.
	CC	The resulting polymer is treated with cyanoogen bromide to remove the Met.
	CC	Copolymer produced in this manner may be used to produce sutures, staples, clips, drug delivery devices, pins and screws.
SQ	Sequence	81 BP; 47 A; 6 C; 4 G; 24 T;
Query Match		Best Local Similarity 0.1%; Score 41; DB I; Length 81; Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	2146	AATGATATATATATATATATATATATATATATATAA
Dd	2	AATGATATATATATATATATATATATATATATATAA
RESULT	4	
ID	I71809	
.AC	G71809 standard; DNA; 103 BP.	
NC	G71809;	
DE	30-MAR-1995 (first entry)	
KW	pSPORIc synthetic polyester gene 5' fragment and trpa initiator.	
KM	Synthetic gene; synthesis; polyester; a-amino acid analogue; clips;	
KW	stop codon; stop signal; monomer unit; lactate; glycolate; staples;	
KX	translation initiation; modified; Met-tRNA ^{fMet} ; cyanoogen bromide; pins;	
KS	co-polymer; sutures; drug delivery device; screws; ds.	
OS	Synthetic.	
FH	Key	Location/Qualifiers
FT	misc_signal	1..18
FI	/tag= a	
FI	/note= "trpa initiator"	
FT	misc_feature	19..21
FT	/tag= b	
FT	/note= "start codon"	
FT	misc_feature	22..93
FT	/tag= c	
FT	/note= "polyester gene 5' fragment"	
PN	CA2112716-A.	
PD	01-JUL-1994.	
PF	31-DEC-1993; 112716.	
PR	31-DEC-1992; US-999520.	
PA	(USDU) US SURGICAL CORP.	
PI	Gruskin EA;	
DR	WPJ; 94-27324/34.	
PT	New biocompatible medical devices and articles - using a lactic acid-glycolic acid polyester in which monomer units are individually and specifically determined	
PS	Example 1; Page 45; 67pp; English.	
CC	This sequence represents a fragment of the plasmid pSPORT ic which contains the 5' fragment of the synthetic genes which encode polyesters,	
CC	given in g71801-02, under the control of the E. coli trpa transcription initiator and followed by PstI and HindIII restriction sites.	
CC	A further fragment of the synthetic genes was ligated into pSPORT ic to give plasmid pSPORT id (see also g71810-11) with the final fragment being inserted in a third round of ligation to give pSPORT iees also g71812-13).	
CC	The E. coli trpa initiator sequence was used to express the full length sequences given in g71801-02 in an E. coli cell free medium.	
CC	The expressed sequences represent synthetic genes which were used in the synthesis of polyesters derived from the a-amino acid analogues of an	


```
FT /mod_base- OTHER
FM /note- OTHER- N6 methyl-8-oxo 2' deoxyadenine"
FT modified_base 344
FT /tag- 1
FT /mod_base- OTHER
FT /note- OTHER- N6 methyl-8-oxo 2' deoxyadenine"
EN W09209705-A.
PD 11-JUN-1992.
PF 25-NOV-1991. U08811.
PR 23-NOV-1990; US-617907.
PR 18-JAN-1991; US-643382.
PR 08-APR-1991; US-683420.
PR 17-APR-1991; US-686544.
PR 17-APR-1991; US-686547.
PR 17-APR-1991; US-686546.
PR 27-SEP-1991; US-766733.
PR (GILE-) GILEAD SCI INC.
PA Firechler B, Krawczyk S, Matteucci MD, Milligan J;
PI WPI: 92-217083/26.
PT New oligomers contg. modified bases - which form a triplex with
PT G-C doublet in a DNA duplex, for treating and diagnosing HIV,
PT hepatitis, herpes, malignancy and inflammation
PS Claim 12, Page 70; 77pp; English.
CC The synthetic oligomer is capable of forming a triplex at
CC physiological pH with a putine rich target sequence by coupling
CC into the major groove of the duplex. The specific target sequence
CC of this oligomer is the human leukocyte adhesion protein P150, 95
CC alpha subunit gene (HUMINTO2) beginning at nucleotide 2370 contg. a
CC putine rich sequence contd. on one strand of the duplex. The oligomer,
CC and others like it are useful in diagnosis and therapy of diseases
CC characterised by specific DNA duplex targets, e.g. HPV; HER; HIV;
CC hepatitis B; herpes, malignant tumours and inflammation. The triple
CC helices form under mild conditions thus assays may be carried out
CC without subjecting the test specimen to harsh conditions. The
CC oligomer contains an inverted polarity region formed from an o-xyloso
CC dimer synthon. The linking sp. is o-xyloso (nucleosides have the 3'
CC positions of xylose sugars linked via the o-xyloso ring). Two
CC nucleotides are coupled through a xylose residue to form the dimer
CC synthon. This additional modification may render the oligomer stable
CC to nuclease activity. The oligomer is able to inhibit gene expression,
CC as verified by in vitro systems.
CC See also Q25452-25501 and Q30226-448.
SQ Sequence 36 BP; 9 A; 0 C; 0 G; 27 T;
```

```
Query Match 0.1%; Score 30; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2158 AATATAATATAATATAATATAATATAA 2187
DB 36 AATATAATATAATATAATATAATATAA 7
```

```
RESULT 8
ID V19044/c
AC V19044;
DE 28-JUL-1998 (first entry)
DE Alu PCR primer 1.
KW PCR; primer; amplification; Alu repeat sequence; vector;
KW circular yeast artificial chromosome; YAC; ss.
OS Synthetic.
OS Saccharomyces sp.
PD 15-JAN-1998.
PR 09-JUL-1996; U11478.
PR 09-JUL-1996; NO-U11478.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA Koudrina NT, Laktionov VL, Perkins EL, Resnick MA;
DR WPI: 98-110234/10.
PT Preparation of yeast artificial chromosomes - by in vivo
PT recombination using vector comprising yeast centromere, marker,
```

```
PT Yeast telomere and nucleic acid for recombination
PS Example 1; Page 45; 117pp; English.
CC This is the nucleotide sequence for the PCR primer used in the
CC amplification of the Alu repeat sequence, which is used to
CC demonstrate the processes described in the invention. It involves
CC the creation and use of circular yeast artificial chromosome (YAC)
CC to selectively clone specific nucleic acids from a background of
CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
CC They can be used to rapidly isolate human DNA where only a part of the
CC sequence of DNA is known. Using the methods large fragments of DNA can
CC be easily cloned and analysed.
SQ Sequence 40 BP; 7 A; 12 C; 13 G; 8 T;
```

```
Query Match 0.1%; Score 30; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 8859 AAGTGTGGGATTACAGCGTGAGCCAC 8888
DB 30 AAGTGTGGGATTACAGCGTGAGCCAC 1
```

RESULT 9

```
ID T26410/c
AC T26410;
DE 06-DEC-1996 (first entry)
DE Human gene signature HUMGS08651.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PD WO9514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUBO/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function. By preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 2078; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
```

```
SQ Sequence 91 BP; 38 A; 22 C; 11 G; 17 T;
```

```
Query Match 0.1%; Score 30; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2460 AGCTGTGCTGTGTCGCCAGGCTGGAGTCC 2489
DB 42 AGCTGTGCTGTGTCGCCAGGCTGGAGTCC 13
```

```
RESULT 10
027391
ID Q27391 standard; DNA; 35 BP.
AC Q27391;
DT 27-JAN-1993 (first entry)
DE Inter-Alu specific primer PDJ33.
KM Polymerase chain reaction; PCR; repetitive element; ss.
OS Synthetic.
PN W09213101-A.
PD 06-AUG-1992.
PF 24-JAN-1992; NL0018.
PR 25-JAN-1991; NL-000132.
PA (INGE-) INGENY BV.
PI Uiterlinden AG, Vljg J;
DR WPI; 92-284683/34.
PT Detection of genetic variation by 2-D electrophoresis of
PT fragments - and hybridisation with labelled probes, carried out
PT on fragments consisting of inter-repeat sequences generated by
PT PCR.
PS Claim 6; Page 6; 31pp; English.
CC Primer PDJ33 is one of several primers which are preferred for use
CC in amplifying inter-Alu regions of DNA. The amplified fragments are
CC then subjected to 2-D electrophoresis on the basis of length and
CC differences in base sequence. The resulting separation pattern is
CC transferred to a filter for screening with a probe. The method can
CC be used to detect genetic variation.
CC See Q27389-Q27404 and Q33141-Q33144.
SQ Sequence 35 BP; 8 A; 10 C; 11 G; 6 T;

Query Match 0.1%; Score 28; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8859 AAAGTCTGGGATTACAGCGGTGAGCCA 8886
DB 8 AAAGTCTGGGATTACAGCGGTGAGCCA 35
|||||
RESULT 11
V83640/C
ID V83640 standard; DNA; 46 BP.
AC V83640;
DT 01-MAR-1999 (first entry)
DE Oligonucleotide 1 used in the construction of assay plasmids.
KM Repetitive sequence; carcinogenic; human dietary component;
KM DNA instability; cancer; diet; primer; ss.
OS Synthetic.
PN W09845476-A1.
PD 15-OCT-1998.
PF 08-APR-1998; G00869.
PR 08-APR-1997; GB-007141.
PA (FOOD-) FOOD RES INST.
PI Schweizer M;
DR WPI; 99-024011/02.
PT Assay for testing the carcinogenic properties of a test substance
PT by introduction of a reporter gene expression vector containing a
PT repetitive DNA sequence that is unstable in cancer cells
PT Disclosure; Page 15; 103pp; English.
CC The present sequence represents an oligonucleotide used in the
CC construction of assay plasmids, which are used in the course of the
CC invention. The specification describes an assay for testing the
CC carcinogenic properties of a test substance. The assay comprises
CC introducing into cells a reporter gene expression vector comprising a
CC repetitive DNA sequence which exhibits instability in cancer cells,
CC whereby instability of the repetitive DNA sequence affects expression of
CC the reporter gene, exposing the resulting cells to the test substance
CC and determining whether the test substance is carcinogenic or
CC anti-carcinogenic by comparing the frequency of reporter gene expression
CC in the resulting cells with the frequency of reporter gene expression in
CC cells which have not been exposed to the test substance. The assay can
CC be used to identify human dietary components that protect against DNA
CC instability, and therefore some types of cancer, and can be used to
```

```
CC contribute to the scientific basis for a healthy diet.
SQ Sequence 46 BP; 18 A; 20 C; 6 G; 2 T;

Query Match 0.1%; Score 28; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18617 GCTGTGTGTGTGTGTGTGTGTGTGTG 18644
DB 40 GCTGTGTGTGTGTGTGTGTGTGTGTGTG 13
|||||
RESULT 12
V83641
ID V83641 standard; DNA; 46 BP.
AC V83641;
DT 01-MAR-1999 (first entry)
DE Oligonucleotide 2 used in the construction of assay plasmids.
KM Repetitive sequence; carcinogenic; human dietary component;
KM DNA instability; cancer; diet; primer; ss.
OS Synthetic.
PN W09845476-A1.
PD 15-OCT-1998.
PF 08-APR-1998; G00869.
PR 08-APR-1997; GB-007141.
PA (FOOD-) FOOD RES INST.
PI Schweizer M;
DR WPI; 99-024011/02.
PT Assay for testing the carcinogenic properties of a test substance
PT by introduction of a reporter gene expression vector containing a
PT repetitive DNA sequence that is unstable in cancer cells
PT Disclosure; Page 16; 103pp; English.
CC The present sequence represents an oligonucleotide used in the
CC construction of assay plasmids, which are used in the course of the
CC invention. The specification describes an assay for testing the
CC carcinogenic properties of a test substance. The assay comprises
CC introducing into cells a reporter gene expression vector comprising a
CC repetitive DNA sequence which exhibits instability in cancer cells,
CC whereby instability of the repetitive DNA sequence affects expression of
CC the reporter gene, exposing the resulting cells to the test substance
CC and determining whether the test substance is carcinogenic or
CC anti-carcinogenic by comparing the frequency of reporter gene expression
CC in the resulting cells with the frequency of reporter gene expression in
CC cells which have not been exposed to the test substance. The assay can
CC be used to identify human dietary components that protect against DNA
CC instability, and therefore some types of cancer, and can be used to
CC contribute to the scientific basis for a healthy diet.
SQ Sequence 46 BP; 2 A; 6 C; 20 G; 18 T;

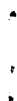
Query Match 0.1%; Score 28; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18617 GCTGTGTGTGTGTGTGTGTGTGTGTG 18644
DB 11 GCTGTGTGTGTGTGTGTGTGTGTGTGTG 38
|||||
RESULT 13
V83635
ID V83635 standard; DNA; 47 BP.
AC V83635;
DT 01-MAR-1999 (first entry)
DE Oligonucleotide 1 used in the construction of assay plasmids.
KM Repetitive oligonucleotide; carcinogenic; human dietary component;
KM DNA instability; cancer; diet; ss.
OS Synthetic.
PN W09845476-A1.
PD 15-OCT-1998.
PF 08-APR-1998; G00869.
PR 08-APR-1997; GB-007141.
```


Mon Apr 3 08:24:29 2000

us-08-852-495c-1_copy_65000_95000.png

Page 7



Mon Apr 3 08:24:30 2000

us-08-852-495c-1_copy_65000_95000.rml

Page 2

```
RESULT 2
US-08-340-426D-69
Sequence 69, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-69

Query Match 0.2%; Score 58; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27685 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGTTCCACCGTGGCCAGAGATG 27742
DB 6 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGTTCCACCGTGGCCAGAGATG 63

RESULT 3
US-08-450-673C-69
Sequence 69, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
COMPUTER TYPE: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2600
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-69

Query Match 0.2%; Score 58; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27685 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGTTCCACCGTGGCCAGAGATG 27742
DB 6 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGTTCCACCGTGGCCAGAGATG 63

RESULT 4
PCT-US95-17111A-69
Sequence 69, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-69
```

Query Match 0.2%; Score 58; DB 6; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27685 GCCCCGCTAATTTTGTATTTTGTAGAGACAGAGGTTTCACCGTGTGGCCAGAG 27742
Db 6 GCCCCGCTAATTTTGTATTTTGTAGAGACAGAGGTTTCACCGTGTGGCCAGAG 63

RESULT 5

US-08-454-557C-57
Sequence 57, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ. ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-57

Query Match 0.2%; Score 55; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27685 GCCCCGCTAATTTTGTATTTTGTAGAGACAGAGGTTTCACCGTGTGGCCAGG 27739
Db 6 GCCCCGCTAATTTTGTATTTTGTAGAGACAGAGGTTTCACCGTGTGGCCAGG 60

RESULT 6

US-08-340-426D-57
Sequence 57, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ. ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-57

Query Match 0.2%; Score 55; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 27685 GCCCCGCTAATTTTGTATTTTGTAGAGACAGAGGTTTCACCGTGTGGCCAGG 27739
Db 6 GCCCCGCTAATTTTGTATTTTGTAGAGACAGAGGTTTCACCGTGTGGCCAGG 60

RESULT 7

US-08-450-673C-57
Sequence 57, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540

Mon Apr 3 08:24:30 2000

us-08-852-495c-1_copy_65000_95000.rni

Page 4

INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673c-57

Query Match 0.2%; Score 55; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27685 GCGCCGCTATTTTGTATTTAGTAGACAGAGGTTTCACCGTTGGCCAGG 27739
DB 6 GCGCCGCTATTTTGTATTTAGTAGACAGAGGTTTCACCGTTGGCCAGG 60

RESULT 8
PCT-US95-17111A-57
Sequence 57, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and
Detection of Alzheimer's Disease
TITLE OF INVENTION: Detection of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-57

Query Match 0.2%; Score 55; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27685 GCGCCGCTATTTTGTATTTAGTAGACAGAGGTTTCACCGTTGGCCAGG 27739
DB 6 GCGCCGCTATTTTGTATTTAGTAGACAGAGGTTTCACCGTTGGCCAGG 60

RESULT 9
US-08-454-557c-92/c

Sequence 92, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
of Alzheimer's Disease
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557c-92

Query Match 0.1%; Score 35; DB 3; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21802 CGGCGGATCAGAGGTGAGGATCGAGACATC 21836
DB 48 CGGCGGATCAGAGGTGAGGATCGAGACATC 14

RESULT 10
US-08-340-426D-92/c
Sequence 92, Application US/08340426D
Patent No. 5948631
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
of Alzheimer's Disease
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D

FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-92

Query Match 0.1%; Score 35; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21802 CGGGCGATCAGAGTCAAGATCGAGACATC 21836
|||||
Db 48 CGGGCGATCAGAGTCAAGATCGAGACATC 14

RESULT 11
US-08-450-673C-92/c
Sequence 92, Application US/08450673C
Patent No. 594888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-92

Query Match 0.1%; Score 35; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21802 CGGGCGATCAGAGTCAAGATCGAGACATC 21836
|||||
Db 48 CGGGCGATCAGAGTCAAGATCGAGACATC 14

RESULT 12
PCT-US95-17111A-92/c
Sequence 92, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-92

Query Match 0.1%; Score 35; DB 6; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21802 CGGGCGATCAGAGTCAAGATCGAGACATC 21836
|||||
Db 48 CGGGCGATCAGAGTCAAGATCGAGACATC 14

RESULT 13
US-08-454-557C-91
Sequence 91, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-91

Query Match 0.1%; Score 35; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-91

Query Match 0.1%; Score 34; DB 3; Length 84;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2676 GCCCGCTCAGCCTCCCAAGTCTGGATTACA 2709
DB 44 GCCCGCTCAGCCTCCCAAGTCTGGATTACA 77

RESULT 14
US-08-340-426D-91
Sequence 91, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid

STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-91

Query Match 0.1%; Score 34; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2676 GCCCGCTCAGCCTCCCAAGTCTGGATTACA 2709
DB 44 GCCCGCTCAGCCTCCCAAGTCTGGATTACA 77

RESULT 15
US-08-450-673C-91
Sequence 91, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-91

Query Match 0.1%; Score 34; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2676 GCCCGCTCAGCCTCCCAAGTCTGGATTACA 2709
DB 44 GCCCGCTCAGCCTCCCAAGTCTGGATTACA 77

Search completed: March 30, 2000, 10:42:50
Job time: 111232 sec

Mon Apr 3 08:24:30 2000

us-08-852-495c-1_copy_65000_95000.rml

Page 7



5
1
4

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2000, 20:38:58 ; Search time 6197.58 Seconds
(without alignments)
13817.943 Million cell updates/sec

Title: US-08-852-495c-1_COPY_65000_95000

Perfect score: 30001

Sequence: 1 TTATTATGACAAATTAAT.....TACAGATACTCTGCAGTTT 30001

Scoring table:

OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database :

EST: *
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_est48:*
83: em_est35:*
84: em_est36:*
85: em_est37:*
86: em_est38:*
87: gb_est49:*
88: gb_est50:*
89: gb_est51:*
90: gb_est52:*
91: gb_est53:*
92: em_est39:*
93: em_est40:*
94: em_est41:*
95: em_est42:*
96: em_est43:*
97: em_est44:*
98: em_est45:*
99: gb_est54:*
100: gb_est55:*
101: em_est46:*
102: gb_est56:*
103: gb_est57:*
104: gb_est58:*
105: gb_est59:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	54	0.2	83	21	R07384	R07384 ye96c03.r1

C	2	53	0.2	91	38	AA780764	AA780767	aa68f12.s
C	3	53	0.2	98	31	AA280198	AA280199	z046412.r
C	4	53	0.2	100	42	A1077628	A1077629	oY26f04.s
C	5	45	0.1	102	100	AQ319270	AQ319271	RPc111-90
C	6	43	0.1	91	50	F29274	F29274	HSPD19063 H
C	7	43	0.1	93	44	A0038446	A0038446	A0038446
C	8	42	0.1	95	34	AA457423	AA457423	aa85b02.r
C	9	41	0.1	73	51	AU077757	AU077757	AU077757
C	10	41	0.1	92	62	A1939667	A1939667	w04f09.x
C	11	41	0.1	104	105	AQ656158	AQ656158	CITR1-E1
C	12	40	0.1	52	20	D25879	D25879	HUMS05672
C	13	40	0.1	97	39	AA907458	AA907458	o102n06.s
C	14	40	0.1	103	81	B48914	B48914	RPc111-4A12
C	15	39	0.1	58	35	C25772	C25772	DcT272
C	16	37	0.1	52	20	D25879	D25879	HUMS05672
C	17	37	0.1	71	35	AA551166	AA551166	NK76d03.s
C	18	37	0.1	90	61	A1864984	A1864984	xW06c11.x
C	19	37	0.1	100	34	AA513537	AA513537	ng71c01.s
C	20	37	0.1	100	34	AA513537	AA513537	ng71c01.s
C	21	37	0.1	105	63	A1977252	A1977252	w319l01.s
C	22	36	0.1	96	50	A1708433	A1708433	ae97c02.x
C	23	36	0.1	105	21	T94466	T94466	yf35b02.r
C	24	35	0.1	35	60	A1801185	A1801185	l079n04.x
C	25	35	0.1	44	36	AA627434	AA627434	nq47907.s
C	26	35	0.1	65	81	B6140	B6140	HS-1038-A1-
C	27	35	0.1	88	38	AA809831	AA809831	oa40f11.s
C	28	35	0.1	96	22	R67088	R67088	y130h05.s1
C	29	35	0.1	100	23	R95968	R95968	yt83e07.r
C	30	35	0.1	102	34	AA481597	AA481597	aa356e1.r
C	31	35	0.1	103	104	AQ582186	AQ582186	RPc1-11-
C	32	35	0.1	105	61	A1838832	A1838832	at72609.x
C	33	34	0.1	71	35	AA557729	AA557729	n147612.s
C	34	34	0.1	84	34	AA501753	AA501753	p050501.s
C	35	34	0.1	86	36	AA601685	AA601685	no02a11.s
C	36	34	0.1	95	35	AA578401	AA578401	n153c01.s
C	37	34	0.1	100	35	AA578614	AA578614	dh23b10.s
C	38	34	0.1	101	30	AA265246	AA265246	z826e09.s
C	39	34	0.1	102	36	AA654562	AA654562	rt75f10.s
C	40	34	0.1	104	35	AA614370	AA614370	np46f03.s
C	41	34	0.1	104	105	AQ656158	AQ656158	CITR1-E1
C	42	34	0.1	105	21	T94466	T94466	yf35b02.r
C	43	34	0.1	105	21	T94466	T94466	yf35b02.r
C	44	33	0.1	68	28	D44637	D44637	HUMS071036
C	45	33	0.1	78	27	AA584630	AA584630	no12e10.s
C	46	33	0.1	83	27	AA019576	AA019576	z612n02.s

ALIGNMENTS

RESULT	1
LOCUS	R07384
DEFINITION	R07384 83 bp mRNA EST 05-APR-1995
ACCESSION	y959c03.r1 Soares fetal liver spleen INFIs Homo sapiens CDNA clone
VERSION	IMAGE:125572.5, similar to gb:u57130_cds1 HISTONE H1D (HUMAN);,
KEYWORDS	mRNA sequence.
SOURCE	R07384 R07384.1 GI:759307
ORGANISM	EST. human. Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Crenata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 83) Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, K., Kucaba, T., Le, M., Lennon, G., Merra, M., Pereaux, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Tevastis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.
TITLE	The Mashu-Merck EST Project.
JOURNAL	Unpublished (1995)
COMMENT	Contact: Wilson RK 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

```

FEATURES
source      location/Qualifiers
            1..83;stem=ucmcc_sarlocase"

Tel.: 314 286 1800
Fax: 314 286 1810
Email: estevatson.wustl.edu
Insert Size: 806
High quality sequence starts: 1 High gally sequence stops: 1
Source: IMAGE Consortium, LBNL This clone is available royalty-free
through LBNL ; contact the IMAGE Consortium (info@image.lbnl.gov)
for further information. Trace considered overall poor quality
Insert length: 806          std Error: 0.00
Seq primer: M13rp1
High quality sequence stop: 1.
location/Qualifiers
1..83;stem=ucmcc_sarlocase"
```

FEATURES

```

/organism="Homo sapiens"
/db_xref="GDB:478117"
/db_xref="taxon:9606"
/clone="IMAGE:125572"
/clone_1lb="Soares fetal liver spleen INFLS"
/sex="male"
/dev_host="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pRT3D (Pharmacia)
with a modified polylinker; Site1: Pac I; Site2: Eco RI;
Site3: Bam HI; Site4: Xba I; Site5: Kpn I; Site6: Sal I;
Site7: Pst I; Site8: Sma I; Site9: Xho I; Site10: Not I;
Site11: Klu I; Site12: Sbf I; Site13: Xba I; Site14: Sma I;
Site15: Kpn I; Site16: Sal I; Site17: Pst I; Site18: Sma I;
Site19: Kpn I; Site20: Sal I; Site21: Pst I; Site22: Sma I;
Site23: Kpn I; Site24: Sal I; Site25: Pst I; Site26: Sma I;
Site27: Kpn I; Site28: Sal I; Site29: Pst I; Site30: Sma I;
Site31: Kpn I; Site32: Sal I; Site33: Pst I; Site34: Sma I;
Site35: Kpn I; Site36: Sal I; Site37: Pst I; Site38: Sma I;
Site39: Kpn I; Site40: Sal I; Site41: Pst I; Site42: Sma I;
Site43: Kpn I; Site44: Sal I; Site45: Pst I; Site46: Sma I;
Site47: Kpn I; Site48: Sal I; Site49: Pst I; Site50: Sma I;
Site51: Kpn I; Site52: Sal I; Site53: Pst I; Site54: Sma I;
Site55: Kpn I; Site56: Sal I; Site57: Pst I; Site58: Sma I;
Site59: Kpn I; Site60: Sal I; Site61: Pst I; Site62: Sma I;
Site63: Kpn I; Site64: Sal I; Site65: Pst I; Site66: Sma I;
Site67: Kpn I; Site68: Sal I; Site69: Pst I; Site70: Sma I;
Site71: Kpn I; Site72: Sal I; Site73: Pst I; Site74: Sma I;
Site75: Kpn I; Site76: Sal I; Site77: Pst I; Site78: Sma I;
Site79: Kpn I; Site80: Sal I; Site81: Pst I; Site82: Sma I;
Site83: Kpn I; Site84: Sal I; Site85: Pst I; Site86: Sma I;
Site87: Kpn I; Site88: Sal I; Site89: Pst I; Site90: Sma I;
Site91: Kpn I; Site92: Sal I; Site93: Pst I; Site94: Sma I;
Site95: Kpn I; Site96: Sal I; Site97: Pst I; Site98: Sma I;
Site99: Kpn I; Site100: Sal I; Site101: Pst I; Site102: Sma I;
Site103: Kpn I; Site104: Sal I; Site105: Pst I; Site106: Sma I;
Site107: Kpn I; Site108: Sal I; Site109: Pst I; Site110: Sma I;
Site111: Kpn I; Site112: Sal I; Site113: Pst I; Site114: Sma I;
Site115: Kpn I; Site116: Sal I; Site117: Pst I; Site118: Sma I;
Site119: Kpn I; Site120: Sal I; Site121: Pst I; Site122: Sma I;
Site123: Kpn I; Site124: Sal I; Site125: Pst I; Site126: Sma I;
Site127: Kpn I; Site128: Sal I; Site129: Pst I; Site130: Sma I;
Site131: Kpn I; Site132: Sal I; Site133: Pst I; Site134: Sma I;
Site135: Kpn I; Site136: Sal I; Site137: Pst I; Site138: Sma I;
Site139: Kpn I; Site140: Sal I; Site141: Pst I; Site142: Sma I;
Site143: Kpn I; Site144: Sal I; Site145: Pst I; Site146: Sma I;
Site147: Kpn I; Site148: Sal I; Site149: Pst I; Site150: Sma I;
Site151: Kpn I; Site152: Sal I; Site153: Pst I; Site154: Sma I;
Site155: Kpn I; Site156: Sal I; Site157: Pst I; Site158: Sma I;
Site159: Kpn I; Site160: Sal I; Site161: Pst I; Site162: Sma I;
Site163: Kpn I; Site164: Sal I; Site165: Pst I; Site166: Sma I;
Site167: Kpn I; Site168: Sal I; Site169: Pst I; Site170: Sma I;
Site171: Kpn I; Site172: Sal I; Site173: Pst I; Site174: Sma I;
Site175: Kpn I; Site176: Sal I; Site177: Pst I; Site178: Sma I;
Site179: Kpn I; Site180: Sal I; Site181: Pst I; Site182: Sma I;
Site183: Kpn I; Site184: Sal I; Site185: Pst I; Site186: Sma I;
Site187: Kpn I; Site188: Sal I; Site189: Pst I; Site190: Sma I;
Site191: Kpn I; Site192: Sal I; Site193: Pst I; Site194: Sma I;
Site195: Kpn I; Site196: Sal I; Site197: Pst I; Site198: Sma I;
Site199: Kpn I; Site200: Sal I; Site201: Pst I; Site202: Sma I;
Site203: Kpn I; Site204: Sal I; Site205: Pst I; Site206: Sma I;
Site207: Kpn I; Site208: Sal I; Site209: Pst I; Site210: Sma I;
Site211: Kpn I; Site212: Sal I; Site213: Pst I; Site214: Sma I;
Site215: Kpn I; Site216: Sal I; Site217: Pst I; Site218: Sma I;
Site219: Kpn I; Site220: Sal I; Site221: Pst I; Site222: Sma I;
Site223: Kpn I; Site224: Sal I; Site225: Pst I; Site226: Sma I;
Site227: Kpn I; Site228: Sal I; Site229: Pst I; Site230: Sma I;
Site231: Kpn I; Site232: Sal I; Site233: Pst I; Site234: Sma I;
Site235: Kpn I; Site236: Sal I; Site237: Pst I; Site238: Sma I;
Site239: Kpn I; Site240: Sal I; Site241: Pst I; Site242: Sma I;
Site243: Kpn I; Site244: Sal I; Site245: Pst I; Site246: Sma I;
Site247: Kpn I; Site248: Sal I; Site249: Pst I; Site250: Sma I;
Site251: Kpn I; Site252: Sal I; Site253: Pst I; Site254: Sma I;
Site255: Kpn I; Site256: Sal I; Site257: Pst I; Site258: Sma I;
Site259: Kpn I; Site260: Sal I; Site261: Pst I; Site262: Sma I;
Site263: Kpn I; Site264: Sal I; Site265: Pst I; Site266: Sma I;
Site267: Kpn I; Site268: Sal I; Site269: Pst I; Site270: Sma I;
Site271: Kpn I; Site272: Sal I; Site273: Pst I; Site274: Sma I;
Site275: Kpn I; Site276: Sal I; Site277: Pst I; Site278: Sma I;
Site279: Kpn I; Site280: Sal I; Site281: Pst I; Site282: Sma I;
Site283: Kpn I; Site284: Sal I; Site285: Pst I; Site286: Sma I;
Site287: Kpn I; Site288: Sal I; Site289: Pst I; Site290: Sma I;
Site291: Kpn I; Site292: Sal I; Site293: Pst I; Site294: Sma I;
Site295: Kpn I; Site296: Sal I; Site297: Pst I; Site298: Sma I;
Site299: Kpn I; Site300: Sal I; Site301: Pst I; Site302: Sma I;
Site303: Kpn I; Site304: Sal I; Site305: Pst I; Site306: Sma I;
Site307: Kpn I; Site308: Sal I; Site309: Pst I; Site310: Sma I;
Site311: Kpn I; Site312: Sal I; Site313: Pst I; Site314: Sma I;
Site315: Kpn I; Site316: Sal I; Site317: Pst I; Site318: Sma I;
Site319: Kpn I; Site320: Sal I; Site321: Pst I; Site322: Sma I;
Site323: Kpn I; Site324: Sal I; Site325: Pst I; Site326: Sma I;
Site327: Kpn I; Site328: Sal I; Site329: Pst I; Site330: Sma I;
Site331: Kpn I; Site332: Sal I; Site333: Pst I; Site334: Sma I;
Site335: Kpn I; Site336: Sal I; Site337: Pst I; Site338: Sma I;
Site339: Kpn I; Site340: Sal I; Site341: Pst I; Site342: Sma I;
Site343: Kpn I; Site344: Sal I; Site345: Pst I; Site346: Sma I;
Site347: Kpn I; Site348: Sal I; Site349: Pst I; Site350: Sma I;
Site351: Kpn I; Site352: Sal I; Site353: Pst I; Site354: Sma I;
Site355: Kpn I; Site356: Sal I; Site357: Pst I; Site358: Sma I;
Site359: Kpn I; Site360: Sal I; Site361: Pst I; Site362: Sma I;
Site363: Kpn I; Site364: Sal I; Site365: Pst I; Site366: Sma I;
Site367: Kpn I; Site368: Sal I; Site369: Pst I; Site370: Sma I;
Site371: Kpn I; Site372: Sal I; Site373: Pst I; Site374: Sma I;
Site375: Kpn I; Site376: Sal I; Site377: Pst I; Site378: Sma I;
Site379: Kpn I; Site380: Sal I; Site381: Pst I; Site382: Sma I;
Site383: Kpn I; Site384: Sal I; Site385: Pst I; Site386: Sma I;
Site387: Kpn I; Site388: Sal I; Site389: Pst I; Site390: Sma I;
Site391: Kpn I; Site392: Sal I; Site393: Pst I; Site394: Sma I;
Site395: Kpn I; Site396: Sal I; Site397: Pst I; Site398: Sma I;
Site399: Kpn I; Site400: Sal I; Site401: Pst I; Site402: Sma I;
Site403: Kpn I; Site404: Sal I; Site405: Pst I; Site406: Sma I;
Site407: Kpn I; Site408: Sal I; Site409: Pst I; Site410: Sma I;
Site411: Kpn I; Site412: Sal I; Site413: Pst I; Site414: Sma I;
Site415: Kpn I; Site416: Sal I; Site417: Pst I; Site418: Sma I;
Site419: Kpn I; Site420: Sal I; Site421: Pst I; Site422: Sma I;
Site423: Kpn I; Site424: Sal I; Site425: Pst I; Site426: Sma I;
Site427: Kpn I; Site428: Sal I; Site429: Pst I; Site430: Sma I;
Site431: Kpn I; Site432: Sal I; Site433: Pst I; Site434: Sma I;
Site435: Kpn I; Site436: Sal I; Site437: Pst I; Site438: Sma I;
Site439: Kpn I; Site440: Sal I; Site441: Pst I; Site442: Sma I;
Site443: Kpn I; Site444: Sal I; Site445: Pst I; Site446: Sma I;
Site447: Kpn I; Site448: Sal I; Site449: Pst I; Site450: Sma I;
Site451: Kpn I; Site452: Sal I; Site453: Pst
```

BASE COUNT
ORIGIN

Query Match	0.28;	Score 54;	DB 21;	Length 83;
Best Local Similarity	100.0%;	Pred. No. 1.6e-06;		
Matches 54;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 13008 CCGCTATGATGTGGAGAAAACACACAGCCGATCAAACTGGTCTCAAGAGCCT 13061
Db 30 CCGCTATGATGTGGAGAAAACACACAGCCGATCAAACTGGTCTCAAGAGCCT 83

RESULT 2

LOCUS	AA780764	91 bp	mRNA	EST	05-FEB-1998
DEFINITION	ac68f12.s1 Strategene fetal retina 937202 Homo sapiens CDNA clone				

ACCESSION	AA780764	
VERSION	AA780764.1	GI:2840095

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;

AUTHORS
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,

TITLE
WashU-NCI human ESR Project
Unpublished (1997)

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

This clone is available royalty-free through JPLI: contact the
 Email: estewatson.wustl.edu
 Fax: 314 286 1810

IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham.

RESULT 5
LOCUS AQ319270 102 bp DNA GSS 06-MAY-1999
DEFINITION RPCI11-98B22.TV Homo sapiens genomic clone RPCI11-98B22,
genomic survey sequence.
ACCESSION AQ319270
VERSION AQ319270.1 GI:4052235
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 102)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE Use of human BAC end Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Other GSSs: RPCI11-98B22.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetlgr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..102
/organism="Homo sapiens"
/db_xref="GDB:7537293"
/db_xref="taxon:9606"
/clone="RPCI11-98B22"
/clone_1lb="RPCI11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: DBAC3.6; Site.1: EcoRI; Site.2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 26 a 28 c 31 g 17 t
ORIGIN
17 t
Query Match 0.1%; Score 45; DB 100; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.00084;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 28844 GGTGATCACCGCTCGGCTCCCAAGTGTGGATTCACAGT 28888
Db 97 GGTGATCACCGCTCGGCTCCCAAGTGTGGATTCACAGT 53
RESULT 6
LOCUS F29274 91 bp mRNA EST 13-MAY-1999
DEFINITION HSPD19063 HM3 Homo sapiens cDNA clone s400074G04, mRNA sequence.
ACCESSION F29274
VERSION F29274.1 GI:4814900
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 91)
AUTHORS Lanfanchi,G., Murro,T., Caldara,F., Pacchioni,B., Pallavicini,A.,
Pandolfo,D., Toppo,S., Trevisan,S., Scarso,S. and Valle,G.

TITLE Identification of 4370 expressed sequence tags from a
3'-end-specific cDNA library of human skeletal muscle by DNA
sequencing and filter hybridization
JOURNAL Genome Res. 6 (1), 35-42 (1996)
MEDLINE 96276048
COMMENT On Apr 7, 1998 this sequence version replaced gi:3034419.
Contact: Valle G.
CIRIBI Biotechnology Centre
University of Padua
Via Trieste 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
http://grup.bio.unipd.it.
FEATURES
source
Location/Qualifiers
1..91
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="s400074G04"
/clone_1lb="HM3"
/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/note="Vector: pCDNAII (Invitrogen); Site.1: BstXI;
Site.2: NotI. This library is not subtracted nor normalized.
Lanfanchi. This library is not subtracted nor normalized.
The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-biotin-AACCGCTGAGTCCAGTACGAGGCGGCTTTTCTTTT-3'). The
ds cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adapters, NotI digested and
directionally cloned into BstXI-NotI cut pCDNAII vector."
BASE COUNT 18 a 19 c 38 g 16 t
ORIGIN
16 t
Query Match 0.1%; Score 43; DB 50; Length 91;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26283 GGGCGCTGAGTCCAGTACGAGGAGGCGTGGAGGAGAA 26325
Db 24 GGGCGCTGAGTCCAGTACGAGGAGGCGTGGAGGAGAA 66
RESULT 7
LOCUS AU038446 93 bp mRNA EST 29-MAR-1999
DEFINITION AU038446 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
discoideum cDNA clone SSH/29, mRNA sequence.
ACCESSION AU038446
VERSION AU038446.1 GI:3985199
KEYWORDS EST.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum
Eukaryota; Dictyostelida; Dictyostelium.
REFERENCE 1 (bases 1 to 93)
AUTHORS Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
Yoshino,R., Mita,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
Williams,J., Maeda,M., Takeuchi,I., Ochial,H. and Tanaka,Y.
TITLE The Dictyostelium developmental cDNA project: generation and
analysis of expressed sequence tags from the first-finger stage of
development
JOURNAL DNA Res. 5 (6), 335-340 (1998)
MEDLINE 99156227
COMMENT On Dec 5, 1997 this sequence version replaced gi:2662913.
Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan
Email: d402huesakura.cc.tsukuba.ac.jp
PROJECT "Dictyostelium discoideum cDNA project in Japan".
FEATURES
source
Location/Qualifiers
1..93

					/organism="Dictyostellium discoideum"
					/strain="AY4"
					/db_xref="taxon:44689"
					/clone="SSH729"
					/clone_1lb="Dictyostellium discoideum SS (H.Drushl哈拉)"
					/dev_stage="slug"
BASE COUNT	57 a	2 c	1 g	33 t	
ORIGIN					
Query Match	0.1%	Score 43:	DB 44:	Length 93:	
Best Local Similarity	100.0%	Pred. No.	0.0037:		
Matches 43:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0	
OY 2146	AATAATATAATATAATATAATATAATATAATATAATATAAT 2188				
Dn 27	AATAATATAATATAATATAATATAATATAATATAATATAAT 69				
RESULT 8					
AA457423/c					
LOCUS	AA457423	95 bp	mRNA	EST	06-JUN-1997
DEFINITION	aab6b02.r1 Stratiogene fetal retina 937202 Homo sapiens cDNA clone				
	IMAGE:838155 5'	similar to contains Alu repetitive element;			
	sequence.				
ACCESSION	AA457423				
VERSION	AA457423.1	GI:2180143			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
	Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
REFERENCE	1 (bases 1 to 95)				
AUTHORS	Hillier,L., Allen,M., Bowles,L., Dubouque,T., Gelsel,G., Jost,S.,				
	Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,				
	Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theisling,B.,				
	White,Y., Wyllie,T., Waterson,R. and Wilson,R.				
TITLE	WashU-Merck EST Project 1997				
JOURNAL	Unpublished (1997)				
COMMENT	On Nov 4, 1996 this sequence version replaced gi:1671271.				
	Contact: Wilson RK				
	Washington University School of Medicine				
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108				
	Tel.: 314 286 1800				
	Fax: 314 286 1810				
	Email: est@watson.wustl.edu				
	This clone is available royalty-free through EMBL; contact the				
	IMAGE Consortium (info@image.llnl.gov) for further information.				
	Seq primer: -28m3 rev1 ET from Amersham				
	High quality sequence stop: 76.				
FEATURES	Location/Qualifiers				
source	1..95				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:838155"				
	/clone_1lb="Stratigene fetal retina 937202"				
	/sex="mixed"				
	/lab_host="SOLR (kanamycin resistant)"				
	/note="Vector: pBluescript SK-; site1: EcoRI; site2:				
	XhoI; cloned unidirectionally. Primer: Oligo dt. Pooled				
	retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR				
	vector; -5' adaptor sequence: 5' GAATTCGGGCGAGCAG 3' -3'				
	adaptor sequence: 5' CTCGAGCTTTTTTTTTTTTTTT 3'"				
BASE COUNT	14 a	39 c	25 g	17 t	
ORIGIN					
Query Match	0.1%	Score 42:	DB 34:	Length 95:	
Best Local Similarity	100.0%	Pred. No.	0.0074:		
Matches 42:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0	
OY 26294	TCGCAGCTACTCGGAGGCTGAGCGCAGGAATGCCGTGTAAC 26335				

DB	84	TCCGAGCTACTCGGAGGCTGAGCAGGAAATGCCTGTAC	43
RESULT	9		
LOCUS	A0072757	73 bp	mRNA EST
DEFINITION	A0072757 Dictyostelium discoidum SS (H.Urushihara)		Dicystostellum
ACCESSION	A0072757		
VERSION	A0072757.1	GI:5179178	
KEYWORDS	EST.		
SOURCE	Dictyostelium discoidum.		
ORGANISM	Eukaryota; Dictyostellida; Dictyostelium.		
REFERENCE	1 (bases 1 to 73)		
AUTHORS	Urushihara,H.		
TITLE	Developmental cDNA in Dictyostelium discoidum (1999)		
JOURNAL	Unpublished (1999)		
COMMENT	On Jun 5, 1998 this sequence version replaced gi:3188495. Contact: Hideko Urushihara Institute of Biological Sciences University of Tsukuba 3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan Email: d402hu@sakura.cc.tsukuba.ac.jp PROJECT - Dictyostelium discoidum cDNA project in Japan.		
FEATURES			
source	1..73		
	/organism="Dictyostelium discoidum"		
	/strain="AX4"		
	/db_xref="taxon:44689"		
	/clone="SSA765"		
	/clone_1lb="Dictyostelium discoidum SS (H.Urushihara)"		
	/dev_stage="slug"		
BASE COUNT	46 a 0 c 1 g 26 t		
ORIGIN			
	Query Match 0.18; Score 41; DB 51; Length 73;		
	Best Local Similarity 100.0%; Pred. No. 0.017;		
	Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	2146 AATAATATATATATATATATATATATATATATATATATTA	2186	
Db	15 AATAATATATATATATATATATATATATATATATATTA	55	
RESULT	10		
LOCUS	A1932967	92 bp	mRNA EST
DEFINITION	wc0f09.x1 NCI-CGAP_Gas4 Homo sapiens CDNA clone IMAGE:2457833		'
ACCESSION	A1932967		
VERSION	A1932967.1	GI:5671704	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
	Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	1 (bases 1 to 92)		
	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
	Tumor Gene Index		
	Unpublished (1997)		
	On Dec 20, 1995 this sequence version replaced gi:1134104.		
	Contact: Robert Strausberg, Ph.D.		
	Tel.: (301) 496-1550		
	Email: Robert.Strausberg@nih.gov		
	Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.		
	Emmert-Buck, M.D., Ph.D.		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: Greg Lennon, Ph.D.		
	DNA Sequencing by: Washington University Genome Sequencing Center		
	Clone distribution: NCI-CGAP clone distribution information can be		

Found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

FEATURES
Seq primer: -40bp from GIPCO.
Location/Qualifiers

SOURCE

1 92
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2457833"
/clone_lib="NCI-CGAP-Gas4"
/tissue_type="Poorly differentiated adenocarcinoma with
signed ring cell features"
/lab_host="DH10B"
/note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: Salt;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.69 kb. Life Technologies catalog #:
11549-011"

BASE COUNT

29 a 19 c 14 g 30 t

ORIGIN

Query Match 0.1%; Score 41; DB 62; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26392 GACAGACGAGACTCCGCTCAAAAAAAAAAAAAAAAA 26432

DB 41 GACAGACGAGACTCCGCTCAAAAAAAAAAAAAAAAA 1

RESULT 11

AA0626158 104 bp DNA GSS 16-JUN-1999
LOCUS CTRBI-E1-2650P18.FR CTRBI-E1 Homo sapiens genomic clone 2650P18,
DEFINITION genomic survey sequence.
ACCESSION AA0626158
VERSION AA0626158.1 GI:5088550
KEYWORDS GSS.

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homindaes; Homo.
1 (bases 1 to 104)
Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building
Unpublished (1997)

JOURNAL

COMMENT Other-GSS: CTRBI-E1-2650P18.TF
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0208
Fax: 301 838 0208

REFERENCE

1 (bases 1 to 104)
Email: hbeetlgr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES

Location/Qualifiers

SOURCE

1. 104
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2650P18"
/clone_lib="CTRBI-E1"
/sex="male"
/cell_type="sperm"
/note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human Library D"

BASE COUNT

12 a 28 c 22 g 42 t

Query Match 0.1%; Score 41; DB 105; Length 104;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26392 GACAGACGAGACTCCGCTCAAAAAAAAAAAAAAAAA 26432

DB 94 GACAGACGAGACTCCGCTCAAAAAAAAAAAAAAAAA 54

RESULT 12

D25879 52 bp mRNA EST 30-NOV-1995
LOCUS HMG505672 Human colon mucosa Homo sapiens CDNA clone cm2333 3',
DEFINITION mRNA sequence.
ACCESSION D25879
VERSION D25879.1 GI:500543
KEYWORDS EST.

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homindaes; Homo.
1 (bases 1 to 52)
Okubo,K., Yoshii,J., Yokouchi,H., Kameyama,M. and Matsubara,K.
Global analysis of gene expression in colon mucosa: a large scale
random cDNA sequencing analysis
Unpublished (1994)

JOURNAL

COMMENT Contact: Okubo,K., Itoh,K., Yoshii,J., Yokouchi,H. and Matsubara,K.
Institute for Molecular and Cellular Biology
Osaka University
3-1 Yamadaoka, Suita, Osaka 565, Japan.
Location/Qualifiers

FEATURES

Source

1. 52
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="cm2333"
/clone_lib="Human colon mucosa"
/note="Adult male, tissue_type = colon mucosa"

BASE COUNT

12 a 17 c 14 g 9 t

ORIGIN

Query Match 0.1%; Score 40; DB 20; Length 52;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26376 GCACTCCAGCGCTGGTGACGAGCGAGACTCCGCTCAA 26415

DB 13 GCACTCCAGCGCTGGTGACGAGCGAGACTCCGCTCAA 52

RESULT 13

AA07458 97 bp mRNA EST 26-AUG-1998
LOCUS AA07458
DEFINITION g102h06.s1 NCI-CGAP_Lus Homo sapiens CDNA clone IMAGE:1522331 3',
similar to contains Alu repetitive element; mRNA sequence.
ACCESSION AA07458
VERSION AA07458.1 GI:3042918
KEYWORDS EST.

ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homindaes; Homo.
1 (bases 1 to 97)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL

COMMENT On Jan 16, 1998 this sequence version replaced g1:1877755.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbp/image/image.html

Insert Length: 624 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 97.
Location/Qualifiers

FEATURES

source

1. .97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1522331"
/clone_1lb="NCI.CGAP.Lu5"
/issue_type="carcinoid"
/lab_host="DH10B"

/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from neuroendocrine lung carcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

17 a 28 c 23 g 27 t 2 others

ORIGIN

Query Match 0.1%; Score 40; DB 39; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8636 TTGACCCAGGCTGAGTGGCGCAATCTCAGCTCA 8675
|||||
Db 5 TTGACCCAGGCTGAGTGGCGCAATCTCAGCTCA 44

RESULT 14

B48914 103 bp DNA GSS 08-APR-1999
LOCUS B48914
DEFINITION RPII11-4A12.TP RPII-11 Homo sapiens genomic clone RPII-11-4A12,
genomic survey sequence.
B48914
B48914.1 GI:2601151

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISMREFERENCE
AUTHORSTITLE
JOURNAL
COMMENT

Human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K.,
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and
Venier, V.C.
Use of BAC End Sequences for Sequence-Ready Map Building
Unpublished (1997)
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPII-11. For BAC
library availability, please contact Pieter de Jong
(pieter@tigr.org, med.bu@tigr.org). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.bu@tigr.org>) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html
Seq primer: Sp6

Class: BAC ends.
Location/Qualifiers

FEATURES

source

1. .103
/organism="Homo sapiens"
/db_xref="GDB:7501163"
/db_xref="taxon:9606"
/clone="RPII-11-4A12"
/clone_1lb="RPII-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;
RPII Human Male BAC Library"

BASE COUNT

30 a 28 c 30 g 15 t

ORIGIN

Query Match 0.1%; Score 40; DB 81; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 26168 TTGGGAGCGCGGCGGATCAGAGCTAGAGAT 26207
|||||
Db 14 TTGGGAGCGCGGCGGATCAGAGCTAGAGAT 53

RESULT 15

C25772 58 bp mRNA EST 29-JUL-1997
LOCUS C25772
DEFINITION Dictyostelium discoideum FC (H. Urushihara) Dictyostelium
discoideum cDNA clone FC-BB11, mRNA sequence.
C25772
C25772.1 GI:2282558

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

EST.
Dictyostelium discoideum.
Dictyostelium discoideum
Eukaryota; Dictyostelida; Dictyostelium.
1 (bases 1 to 58)
Sexual cDNA in D. discoideum(970724)
Suzuki, K., Shimizu, H. and Urushihara, H.
Unpublished (1997)
On Apr 14, 1993 this sequence version replaced g1:693374.
Contact: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
Email: d402nuesakura.cc.tsukuba.ac.jp.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Location/Qualifiers
1. .58
/organism="Dictyostelium discoideum"
/strain="KAX3"
/db_xref="taxon:44689"
/clone="FC-BB11"
/clone_1lb="Dictyostelium discoideum FC (H. Urushihara)"

FEATURES

source

BASE COUNT 35 a 6 c 1 g 16 t
ORIGIN

Query Match 0.1%; Score 39; DB 35; Length 58;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2147 ATATATATATATATATATATATATATATATATATA 2165
|||||
Db 1 ATATATATATATATATATATATATATATATATATA 39

Search completed: March 30, 2000, 03:49:55
Job time: 88449 sec

Mon Apr 3 08:24:30 2000

us-08-852-495c-1_copy_65000_95000.rst

Page 8

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 07:33:31 ; Search time 13412.7 Seconds

(without alignments) updates/sec

Title: US-08-852-495c-1_copy_90000_120000
Perfect score: 30001
Sequence: 1 TGGCCAGCTGCTGAGTGGCGG.....TATTCATAGCCATGTCATAT 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 821193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_com:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pr2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_v1:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_com:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_v1:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*

44: gb_htg6:*
45: em_htg1:*
46: em_htg2:*
47: em_htg3:*
48: em_htg4:*
49: em_hum5:*
50: gb_pl3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	0.2	76	5 AR051499	AR051499 Sequence
2	55	0.2	60	5 AR051487	AR051487 Sequence
3	48	0.2	48	5 AR032547	AR032547 Sequence
4	48	0.2	48	5 I29287	I29287 Sequence 15
5	48	0.2	48	5 I29287	I29287 Sequence 15
6	47	0.2	47	5 AR032558	AR032558 Sequence
7	47	0.2	47	5 I29298	I29298 Sequence 17
8	47	0.2	47	5 I29298	I29298 Sequence 17
9	47	0.2	47	5 I29298	I29298 Sequence 17
10	46	0.2	46	5 I29298	I29298 Sequence 16
11	46	0.2	46	5 I29298	I29298 Sequence 16
12	45	0.1	45	5 AR032557	AR032557 Sequence
13	45	0.1	45	5 I29297	I29297 Sequence 16
14	45	0.1	45	5 I29297	I29297 Sequence 16
15	37	0.1	79	10 S73203	S73203 ATL-1 (rand
16	36	0.1	66	40 AF087511	AF087511 Homo sapi
17	35	0.1	57	13 HUMOT1543A	HUMOT1543A
18	34	0.1	42	9 HUMALUANC	HUMALUANC
19	33	0.1	65	9 HUMALUANC	HUMALUANC
20	32	0.1	63	3 DOGSR7SL1	DOGSR7SL1
21	32	0.1	80	9 HUMBRKFC	HUMBRKFC
22	31	0.1	66	10 S54071	S54071
23	31	0.1	67	13 HUMOT44A	HUMOT44A
24	31	0.1	69	13 HUMOT57BA	HUMOT57BA
25	31	0.1	90	9 HUMDLRFL	HUMDLRFL
26	31	0.1	100	11 HSU67848	HSU67848 Human beta-
27	31	0.1	101	13 HUMOT832A	HUMOT832A
28	31	0.1	104	5 A08899	A08899 H. sapiens (
29	30	0.1	30	5 A38240	A38240 Sequence 4
30	30	0.1	30	5 A38241	A38241 Sequence 5
31	30	0.1	30	5 A38242	A38242 Sequence 6
32	30	0.1	66	40 AF087511	AF087511 Homo sapi
33	30	0.1	67	13 HUMOT44A	HUMOT44A
34	30	0.1	73	5 A08915	A08915 H. sapiens (
35	30	0.1	93	5 HSHTPKIB	HSHTPKIB
36	30	0.1	100	9 HSLSA27	HSLSA27
37	30	0.1	102	5 A08911	A08911 H. sapiens (
38	30	0.1	105	13 HS015Xa3	HS015Xa3
39	29	0.1	40	5 A68621	A68621 Sequence 1
40	29	0.1	65	9 HUMALUANC	HUMALUANC
41	29	0.1	75	11 AF032333	AF032333 Propithec
42	29	0.1	76	11 AF032287	AF032287 Eulemur m
43	29	0.1	88	13 HUMOT5056B	HUMOT5056B
44	29	0.1	99	5 A08898	A08898 H. sapiens (
45	29	0.1	100	9 HUMGALNSA	HUMGALNSA

ALIGNMENTS

RESULT 1
LOCUS AR051499 76 bp DNA
DEFINITION AR051499 Sequence 69 from patent US 5830670.
ACCESSION AR051499
VERSION AR051499.1 GI:5974863

29-SEP-1999

Mon Apr 3 08:24:31 2000

us-08-852-495c-1_copy_90000_120000.1rge

Page 2

KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 76)
AUTHORS de la Monte,S. and Wands,J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 69 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..76
BASE COUNT 15 a 19 c 18 g 24 t
ORIGIN

Query Match 0.2%; Score 58; DB 5; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.1e-18;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2685 GCGCCGCTAATTTTGTATTTTGTAGTACAGACAGGGTTTCACCGTGTGGCCAGGATG 2742
Db 6 GCGCCGCTAATTTTGTATTTTGTAGTACAGACAGGGTTTCACCGTGTGGCCAGGATG 63

RESULT 2
LOCUS AR051487 60 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 57 from patent US 5830670.
ACCESSION AR051487
VERSION AR051487.1 GI:5974851
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 60)
TITLE de la Monte,S. and Wands,J.R.
JOURNAL Patent: US 5830670-A 57 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..60
BASE COUNT 12 a 14 c 15 g 19 t
ORIGIN

Query Match 0.2%; Score 55; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 4.1e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2685 GCGCCGCTAATTTTGTATTTTGTAGTACAGACAGGGTTTCACCGTGTGGCCAGG 2739
Db 6 GCGCCGCTAATTTTGTATTTTGTAGTACAGACAGGGTTTCACCGTGTGGCCAGG 60

RESULT 3
LOCUS AR032547 48 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 159 from patent US 5869241.
ACCESSION AR032547
VERSION AR032547.1 GI:5948152
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 48)
TITLE Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
JOURNAL Method of determining DNA sequence preference of a DNA-binding molecule
Patent: US 5869241-A 159 09-FEB-1999;
FEATURES Location/Qualifiers
source 1..48

BASE COUNT 12 a 8 c 14 g 14 t
ORIGIN

Query Match 0.2%; Score 48; DB 5; Length 48;
Best Local Similarity 100.0%; Pred. No. 2e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 22582 GAGAGAAATAGGAACCGATCTTTATATATCTACGTTACCCCTGCCCCC 22629
Db 48 GAGAGAAATAGGAACCGATCTTTATATATCTACGTTACCCCTGCCCCC 1

RESULT 4
LOCUS 129287 48 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 159 from patent US 5578444.
ACCESSION 129287
VERSION 129287.1 GI:1820078
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 48)
TITLE Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
JOURNAL Sequence-directed DNA-binding molecules compositions and methods
Patent: US 5578444-A 159 26-NOV-1996;
FEATURES Location/Qualifiers
source 1..48
BASE COUNT 12 a 8 c 14 g 14 t
ORIGIN

Query Match 0.2%; Score 48; DB 5; Length 48;
Best Local Similarity 100.0%; Pred. No. 2e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 22582 GAGAGAAATAGGAACCGATCTTTATATATCTACGTTACCCCTGCCCCC 22629
Db 48 GAGAGAAATAGGAACCGATCTTTATATATCTACGTTACCCCTGCCCCC 1

RESULT 5
LOCUS 190961 48 bp DNA PAT 01-DEC-1998
DEFINITION Sequence 159 from patent US 5726014.
ACCESSION 190961
VERSION 190961.1 GI:3935431
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 48)
TITLE Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.
JOURNAL Screening assay for the detection of DNA-binding molecules
Patent: US 5726014-A 159 10-MAR-1998;
FEATURES Location/Qualifiers
source 1..48
BASE COUNT 12 a 8 c 14 g 14 t
ORIGIN

RESULT 6
AR032558 47 bp DNA PAT 29-SEP-1999
LOCUS AR032558
DEFINITION Sequence 170 from patent US 5869241.
ACCESSION AR032558
VERSION AR032558.1 GI:5948163
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 47)
TITLE Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
METHOD of determining DNA sequence preference of a DNA-binding molecule
JOURNAL Patent: US 5869241-A 170 09-FEB-1999;
FEATURES Location/Qualifiers
source 1..47
BASE COUNT 13 a 8 c 10 g 16 t
ORIGIN

Query Match 0.2%; Score 47; DB 5; Length 47;
Best Local Similarity 100.0%; Pred. No. 6.6e-13; Mismatches 0; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12078 CAACTAGTGGTGTGCTGACTCTATAAAGAGAAGTAGCTCTTCCTT 12124
|||||
Db 1 CAACTAGTGGTGTGCTGACTCTATAAAGAGAAGTAGCTCTTCCTT 47

RESULT 7
129298 47 bp DNA PAT 06-FEB-1997
LOCUS 129298
DEFINITION Sequence 170 from patent US 5578444.
ACCESSION 129298
VERSION 129298.1 GI:1820089
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 47)
TITLE Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
METHOD of determining DNA sequence preference of a DNA-binding molecule
JOURNAL Patent: US 5578444-A 170 26-NOV-1996;
FEATURES Location/Qualifiers
source 1..47
BASE COUNT 13 a 8 c 10 g 16 t
ORIGIN

Query Match 0.2%; Score 47; DB 5; Length 47;
Best Local Similarity 100.0%; Pred. No. 6.6e-13; Mismatches 0; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12078 CAACTAGTGGTGTGCTGACTCTATAAAGAGAAGTAGCTCTTCCTT 12124
|||||
Db 1 CAACTAGTGGTGTGCTGACTCTATAAAGAGAAGTAGCTCTTCCTT 47

RESULT 8
190972 47 bp DNA PAT 01-DEC-1998
LOCUS 190972
DEFINITION Sequence 170 from patent US 5726014.
ACCESSION 190972
VERSION 190972.1 GI:3935442
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 47)
TITLE Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
METHOD of determining DNA sequence preference of a DNA-binding molecule
JOURNAL Patent: US 5726014-A 170 10-MAR-1998;
FEATURES Location/Qualifiers
source 1..47
BASE COUNT 10 a 16 c 6 g 14 t
ORIGIN

Query Match 0.2%; Score 46; DB 5; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.2e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10562 CACAGCCTACCTCCAGTCAGTATAAATCTCTGCTGCGGTC 10607
|||||
Db 1 CACAGCCTACCTCCAGTCAGTATAAATCTCTGCTGCGGTC 46

RESULT 10
129295 46 bp DNA PAT 06-FEB-1997
LOCUS 129295
DEFINITION Sequence 167 from patent US 5578444.
ACCESSION 129295
VERSION 129295.1 GI:1820086
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 46)
TITLE Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
METHOD of determining DNA sequence preference of a DNA-binding molecule
JOURNAL Patent: US 5578444-A 167 26-NOV-1996;
FEATURES Location/Qualifiers
source 1..46
BASE COUNT 10 a 16 c 6 g 14 t
ORIGIN

Query Match 0.2%; Score 46; DB 5; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.2e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10562 CACAGCCTACCTCCAGTCAGTATAAATCTCTGCTGCGGTC 10607
|||||
Db 1 CACAGCCTACCTCCAGTCAGTATAAATCTCTGCTGCGGTC 46

AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.
TITLE Screening assay for the detection of DNA-binding molecules
JOURNAL Patent: US 5726014-A 170 10-MAR-1998;
FEATURES Location/Qualifiers
source 1..47
BASE COUNT 13 a 8 c 10 g 16 t
ORIGIN

Query Match 0.2%; Score 47; DB 5; Length 47;
Best Local Similarity 100.0%; Pred. No. 6.6e-13; Mismatches 0; Indels 0; Gaps 0;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12078 CAACTAGTGGTGTGCTGACTCTATAAAGAGAAGTAGCTCTTCCTT 12124
|||||
Db 1 CAACTAGTGGTGTGCTGACTCTATAAAGAGAAGTAGCTCTTCCTT 47

RESULT 9
AR032555 46 bp DNA PAT 29-SEP-1999
LOCUS AR032555
DEFINITION Sequence 167 from patent US 5869241.
ACCESSION AR032555
VERSION AR032555.1 GI:5948160
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 46)
TITLE Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
METHOD of determining DNA sequence preference of a DNA-binding molecule
JOURNAL Patent: US 5869241-A 167 09-FEB-1999;
FEATURES Location/Qualifiers
source 1..46
BASE COUNT 10 a 16 c 6 g 14 t
ORIGIN

Query Match 0.2%; Score 46; DB 5; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.2e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10562 CACAGCCTACCTCCAGTCAGTATAAATCTCTGCTGCGGTC 10607
|||||
Db 1 CACAGCCTACCTCCAGTCAGTATAAATCTCTGCTGCGGTC 46

RESULT 9
AR032555 46 bp DNA PAT 29-SEP-1999
LOCUS AR032555
DEFINITION Sequence 167 from patent US 5869241.
ACCESSION AR032555
VERSION AR032555.1 GI:5948160
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 46)
TITLE Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
METHOD of determining DNA sequence preference of a DNA-binding molecule
JOURNAL Patent: US 5869241-A 167 09-FEB-1999;
FEATURES Location/Qualifiers
source 1..46
BASE COUNT 10 a 16 c 6 g 14 t
ORIGIN

Query Match 0.2%; Score 46; DB 5; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.2e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10562 CACAGCCTACCTCCAGTCAGTATAAATCTCTGCTGCGGTC 10607
|||||
Db 1 CACAGCCTACCTCCAGTCAGTATAAATCTCTGCTGCGGTC 46

RESULT 10
129295 46 bp DNA PAT 06-FEB-1997
LOCUS 129295
DEFINITION Sequence 167 from patent US 5578444.
ACCESSION 129295
VERSION 129295.1 GI:1820086
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 46)
TITLE Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
METHOD of determining DNA sequence preference of a DNA-binding molecule
JOURNAL Patent: US 5578444-A 167 26-NOV-1996;
FEATURES Location/Qualifiers
source 1..46
BASE COUNT 10 a 16 c 6 g 14 t
ORIGIN

Query Match 0.2%; Score 46; DB 5; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.2e-12; Mismatches 0; Indels 0; Gaps 0;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10562 CACAGCCTACCTCCAGTCAGTATAAATCTCTGCTGCGGTC 10607
|||||
Db 1 CACAGCCTACCTCCAGTCAGTATAAATCTCTGCTGCGGTC 46

```
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 10562 CACAGCCTACCTCCAGTCAGTAAATACCTCTCGCTGCGTTC 10607
Db 1 CACAGCCTACCTCCAGTCAGTAAATACCTCTCTCGCTGCGTTC 46

RESULT 11
LOCUS 190969 46 bp DNA PAT 01-DEC-1998
DEFINITION Sequence 167 from patent US 5726014.
ACCESSION 190969
VERSION 190969.1 GI:3935439
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 46)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.
TITLE Screening assay for the detection of DNA-binding molecules
JOURNAL Patent: US 5726014-A 167 10-MAR-1998;
FEATURES
1.46
Location/Qualifiers
BASE COUNT 10 a 16 c 6 g 14 t
ORIGIN

Query Match 0.2%; Score 46; DB 5; Length 46;
Best Local Similarity 100.0%; Fred. No. 2.2e-12;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 10562 CACAGCCTACCTCCAGTCAGTAAATACCTCTCTCGCTGCGTTC 10607
Db 1 CACAGCCTACCTCCAGTCAGTAAATACCTCTCTCGCTGCGTTC 46

RESULT 12
LOCUS AR032557 45 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 169 from patent US 5869241.
ACCESSION AR032557
VERSION AR032557.1 GI:5948162
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 45)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
TITLE Method of determining DNA sequence preference of a DNA-binding molecule
JOURNAL Patent: US 5869241-A 169 09-FEB-1999;
FEATURES
1.45
Location/Qualifiers
BASE COUNT 8 a 6 c 10 g 21 t
ORIGIN

Query Match 0.1%; Score 45; DB 5; Length 45;
Best Local Similarity 100.0%; Fred. No. 7.4e-12;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12057 ATTTTGAATTTCTTGCGTCAATAGTGTGCTGACTCTAT 12101
Db 1 ATTTTGAATTTCTTGCGTCAATAGTGTGCTGACTCTAT 45

RESULT 13
LOCUS 129297 45 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 169 from patent US 5578444.
ACCESSION 129297
```

```
VERSION 129297.1 GI:1820088
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 45)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
TITLE Sequence directed DNA-binding molecules compositions and methods
JOURNAL Patent: US 5578444-A 169 26-NOV-1996;
FEATURES
1.45
Location/Qualifiers
BASE COUNT 8 a 6 c 10 g 21 t
ORIGIN

Query Match 0.1%; Score 45; DB 5; Length 45;
Best Local Similarity 100.0%; Fred. No. 7.4e-12;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12057 ATTTTGAATTTCTTGCGTCAATAGTGTGCTGACTCTAT 12101
Db 1 ATTTTGAATTTCTTGCGTCAATAGTGTGCTGACTCTAT 45

RESULT 14
LOCUS 190971 45 bp DNA PAT 01-DEC-1998
DEFINITION Sequence 169 from patent US 5726014.
ACCESSION 190971
VERSION 190971.1 GI:3935441
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 45)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.
TITLE Screening assay for the detection of DNA-binding molecules
JOURNAL Patent: US 5726014-A 169 10-MAR-1998;
FEATURES
1.45
Location/Qualifiers
BASE COUNT 8 a 6 c 10 g 21 t
ORIGIN

Query Match 0.1%; Score 45; DB 5; Length 45;
Best Local Similarity 100.0%; Fred. No. 7.4e-12;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12057 ATTTTGAATTTCTTGCGTCAATAGTGTGCTGACTCTAT 12101
Db 1 ATTTTGAATTTCTTGCGTCAATAGTGTGCTGACTCTAT 45

RESULT 15
LOCUS S73203 79 bp DNA PRI 28-FEB-1995
DEFINITION ALU-1 (tandem duplication) [human, acute myeloid leukemia patient,
ACCESSION S73203
VERSION S73203.1 GI:685048
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 79)
AUTHORS Schichman,S.A., Caligiuri,M.A., Strout,M.P., Carter,S.L., Gu,Y.,
Canani,R., Bloomfield,C.D. and Croce,C.M.
TITLE ALU-1 tandem duplication in acute myeloid leukemia with a normal
karyotype involves homologous recombination between Alu elements
Cancer Res. 54 (16), 4277-4280 (1994)
```

MEDLINE 94320053
 GenBank staff at the National Library of Medicine created this
 entry [NCBI gidsq 15288] from the original journal article.

This sequence comes from Fig. 3a.
 Map location: 11.

FEATURES
 source

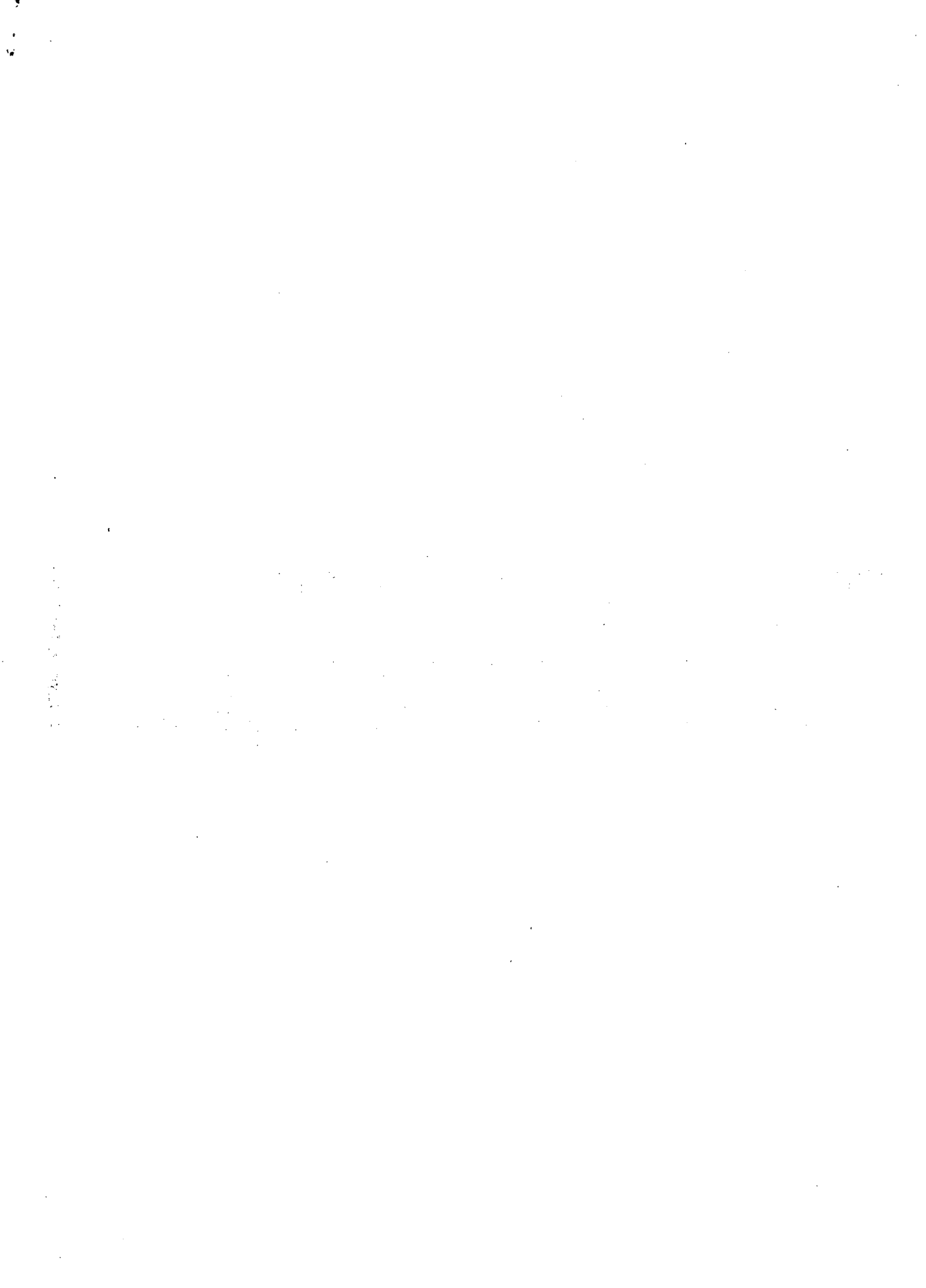
1. .79
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1. .79
 /partial
 /gene="ALT-1"

BASE COUNT 19 a 18 c 29 g 13 t
 ORIGIN

Query Match 0.1%; Score 37; DB 10; Length 79;
 Best Local Similarity 100.0%; Pred. No. 1.2e-07;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5166 TCTCAGCTCACTGCACCTCGCCTCCTGGGTCAAG 5202
 ||||||||||||||||||||||||||||||||||||
 Db 74 TCTCAGCTCACTGCACCTCGCCTCCTGGGTCAAG 38

Search completed: March 30, 2000, 17:03:40
 Job time: 135569 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 10:53:10 ; Search time 611.3 Seconds
(without alignments)
12278.771 Million cell updates/sec

Title: US-08-852-495c-1_COPY_90000_120000

Perfect score: 30001
Sequence: 1 TGGCCAGCTGCGAGTGGCGG.....TATTCAGCATGTCAT 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 311585 seqs, 125096042 residues

Word size : 0

Total number of hits satisfying chosen parameters: 433264

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	48	0.2	48	069409	Human H4/a gene fo
2	48	0.2	48	T63871	Human H4/a gene (f
3	48	0.2	48	T17159	Test sequence from
4	47	0.2	47	069420	Human H3 histone g
5	47	0.2	47	T63882	Human H3 histone g
6	47	0.2	47	T17170	Test sequence from
7	46	0.2	46	069417	Human H2a histone
8	46	0.2	46	T63879	Human H2a histone
9	45	0.2	45	T17167	Test sequence from
10	45	0.1	45	069419	Human H3 histone g
11	45	0.1	45	T63881	Human H3 histone g
12	45	0.1	45	T17169	Test sequence from
13	40	0.1	30	062216	Histone H3 mRNA pr
14	30	0.1	30	062217	Histone H3 mRNA pr
15	30	0.1	30	062218	Histone H3 mRNA pr
16	29	0.1	40	V19044	Alu PCR primer 1.
17	28	0.1	60	X19944	OLIGONUCLEOTIDE SE
18	28	0.1	63	T33701	Microsatellite seq
19	28	0.1	68	T23904	Human gene signatu
20	27	0.1	36	V09783	Tobacco PABF probe
21	27	0.1	44	033855	Microsatellite seq
22	27	0.1	44	033855	Microsatellite seq
23	27	0.1	57	T65739	Repeat sequence fr
24	27	0.1	62	034053	Microsatellite seq
25	27	0.1	70	034155	Sequence of a micr
26	26	0.1	32	027389	Inter-Alu specific
27	26	0.1	35	027391	Inter-Alu specific
28	26	0.1	36	V09783	Tobacco PABF probe
29	26	0.1	42	086155	Sindbis polyA prim
30	26	0.1	42	T30789	Sindbis genomic cd
31	26	0.1	42	T35054	Sindbis virus stra
32	26	0.1	42	V42366	PCR primer 4B used
33	26	0.1	42	V60127	PCR primer 4B used

34	26	0.1	42	1	V70686	Primer 4B used to
35	26	0.1	48	1	086183	Primer SINdball700
36	26	0.1	48	1	086183	Primer SINdball700
37	26	0.1	48	1	T30807	Sindbis PCR primer
38	26	0.1	48	1	T30807	Sindbis PCR primer
39	26	0.1	48	1	T35073	Sindbis-based, tum
40	26	0.1	48	1	T35073	Sindbis-based, tum
41	26	0.1	48	1	V42384	Reverse PCR primer
42	26	0.1	48	1	V42384	Reverse PCR primer
43	26	0.1	48	1	V70704	Reverse PCR primer
44	26	0.1	48	1	V70704	Reverse PCR primer
45	26	0.1	53	1	Q33621	Microsatellite seq

ALIGNMENTS

RESULT 1	
ID 069409/C	069409 standard; DNA; 48 BP.
AC 069409;	
DT 27-FEB-1995 (first entry)	
DE Human H4/a gene for H4 histone, target region.	
KM DNA protein-binding assay: test sequence; screening sequence;	
KW promoter: target; TATA box; Herpes Simplex Virus; HSV;	
KW origin of replication; UL5; transcription factor; TRID: ds.	
OS Synthetic.	
PN W0941980-A.	
PD 07-JUL-1994.	
PE 20-DEC-1993; U12388.	
PR 23-DEC-1993; US-996783.	
PR 17-SEP-1993; US-123936.	
PA (GENE-) GENELABS TECHNOLOGIES INC.	
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;	
DR WPI: 94-234711/28.	
PT Sequence-directed DNA-binding molecules - useful in	
PS pharmaceuticals and as molecular reagents	
PT Claim 28: Page 291: 587pp: English.	
CC A DNA protein-binding assay is provided, useful for screening	
CC libraries of synthetic or biological cpds. for their ability	
CC to bind DNA test sequences. The assay is versatile in that any	
CC number of test sequences can be tested by placing the test sequence	
CC adjacent to a defined protein-binding screening sequence. Binding	
CC of mols. to these test sequences changes the binding characteristics	
CC of the protein mol. to its cognate binding sequence. When such a mol.	
CC binds the test sequence, the equilibrium of the DNA:protein complexes	
CC is disturbed, generating changes in the concentration of free DNA probe.	
CC One application of this method is to eucaryotic general transcription	
CC factors (e.g. TFIID), where the target region is typically selected	
CC from DNA sequences adjacent to the binding site for the eucaryotic	
CC transcription factor. Numerous exemplary test sequences are given:	
CC the sequences in Q69251-731 and Q69550 correspond to promoter targets	
CC (typically, TATA box-contg. sites) for human genes and the sequences in	
CC Q69733-849 correspond to promoter targets for viral genes. The test	
CC sequences may also be randomly generated. DNA:protein interaction may	
CC be used for screening purposes, e.g. the Herpes Simplex Virus (HSV)	
CC origin of replication and UL5 (see Q69551-52, Q69865 and Q69891).	
CC Sequence 48 BP; 12 A; 8 C; 14 G; 14 T;	
SO	
Query Match	0.2%; Score 48; DB 1; Length 48;
Best Local Similarity	100.0%; Pred. No. 9.7e-05;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 22582 GAGAGATAGCAACCGATCTTATATATACGTCACCCCTGCCGCC 22629	
DB 48 GAGAGATAGCAACCGATCTTATATATACGTCACCCCTGCCGCC 1	
RESULT 2	
ID T63871/C	T63871 standard; DNA; 48 BP.
AC T63871;	

DT 14-MAR-1997 (first entry)
DE Human H4/a gene (for Histone H4) TFIIID binding site.
KW Duplex DNA; target region; binding characteristic; DNA binding protein;
KW TFIIID; transcription factor; binding site; inhibition; enhance;
KW Cancer; inherited genetic disorder; alpha-D-galactosidase A; ds.
OS Homo sapiens.
PN US557844.2.
PD 26-NOV-1996.
PF 27-JUN-1991; 723618.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PR 20-DEC-1993; US-171389.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
DR WPI: 97-020402/02.
PT Altering binding characteristics of DNA binding proteins to duplex
PT DNA - by attaching specific small cpd. to target region close to the
PT protein's binding site, useful in treatment of viral disease, cancer
PT etc.
PS Claim 6: Column 179-180; 264pp; English.
CC The sequences given in T63713-4312 represent duplex DNA's which act
CC as target regions in the method of the invention. The method for
CC altering the binding characteristics of a DNA-binding protein to duplex
CC DNA comprises contacting the duplex DNA with a small molecule which
CC binds sequence-specifically to a target region, where, when the small
CC molecule is bound to the target region, it is adjacent to, but not
CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
CC The small molecule is added at a concentration effective to alter the
CC binding of the DNA binding protein. The small molecule may inhibit or enhance
CC the binding of the DNA-binding protein to its binding site. The
CC compounds isolated using this method are potentially useful as
CC therapeutic agents for treatment of any disease which involves a
CC specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.
CC The method is suitable for screening large biological or chemical
CC libraries and allows determination of sequence-specific and relative
CC affinities of known DNA-binding agents for different DNA sequences.
CC The design of these duplex DNA's allows a single DNA:protein interaction
CC to be used for screening sequence-specific, or preferential, DNA binding
CC proteins that recognise almost any possible sequence (see also T49539-
CC 74).
SQ Sequence 48 BP; 12 A; 8 C; 14 G; 14 T;
Query Match 0.2%; Score 48; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 9.7e-05;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 22582 GAGAGATGGAACCGATCTTATATCTAGCTTACCCGCCGCC 22629
DB 48 GAGAGATGGAACCGATCTTATATCTAGCTTACCCGCCGCC 1
RESULT 3
X17159/c
ID X17159 standard; DNA: 48 BP.
AC X17159;
DT 06-MAY-1999 (first entry)
DE Test sequence from human H4/a gene for H4 histone.
KW Test sequence; DNA-binding molecule; screening sequence; human;
KW nucleic acid amplification; target; viral; ds.
OS Homo sapiens.
PN US5869241-A.
PD 09-FEB-1999.
PF 07-JUN-1995; 475228.
PR 20-DEC-1993; US-171389.
PR 27-JUN-1991; US-723618.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PR 07-JUN-1995; US-475228.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;

DR WPI: 99-152755/13.
PT Determination of DNA sequence preference of a DNA-binding molecule -
PT based on inhibition of binding of protein to oligonucleotide
PT sequence attached to test sequence
PS Claim 3: Columns 181-182; 270pp; English.
CC Sequences X17001 to X17600 represent specifically claimed target test
CC sequences that are used in the method of the invention of determining the
CC DNA sequence preference of a DNA-binding molecule. The method comprises:
CC (i) adding a test molecule and a DNA-binding protein to a mixture of
CC duplex DNA test oligonucleotides, each of the test oligonucleotides
CC having a test sequence adjacent to a screening sequence, where the
CC screening sequence binds to the DNA-binding protein with a binding
CC affinity that is independent of the DNA sequence of the test sequence,
CC and where the mixture of duplex DNA test oligonucleotides includes
CC several test sequences; (ii) incubating the test molecule, the mixture of
CC duplex DNA test oligonucleotides and the DNA-binding protein for a time
CC sufficient to permit binding of the test molecule to test sequences in
CC the duplex DNA; (iii) separating unbound test oligonucleotides from test
CC oligonucleotides bound to binding protein; (iv) amplifying the unbound
CC test oligonucleotides; (v) repeating steps (ii) to (iv); (vi) isolating
CC the amplified test oligonucleotides; and (vii) sequencing the isolated
CC test oligonucleotides. Test sequences X17001-X17481 and X17600 correspond
CC to promoter targets for human genes and test sequences X17482-X17599
CC correspond to promoter targets for viral genes.
SQ Sequence 48 BP; 12 A; 8 C; 14 G; 14 T;
Query Match 0.2%; Score 48; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 9.7e-05;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 22582 GAGAGATGGAACCGATCTTATATCTAGCTTACCCGCCGCC 22629
DB 48 GAGAGATGGAACCGATCTTATATCTAGCTTACCCGCCGCC 1
RESULT 4
ID 069420
AC 069420;
DT 27-FEB-1995 (first entry)
DE Human H3 histone gene, target region.
KW DNA protein-binding assay; test sequence; screening sequence;
KW promoter; target; TATA box; Herpes Simplex Virus; HSV;
KW origin of replication; UL9; transcription factor; TFIIID; ds.
OS Synthetic.
PN WO9414980-A.
PD 07-JUL-1994.
PF 20-DEC-1993; U12388.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
DR WPI: 94-234711/28.
PT Sequence-directed DNA-binding molecules - useful in
PT pharmaceuticals and as molecular reagents
PS Claim 26: Page 29; 587pp; English.
CC A DNA protein-binding assay is provided, useful for screening
CC libraries of synthetic or biological cpds. for their ability
CC to bind DNA test sequences. The assay is versatile in that any
CC number of test sequences can be tested by placing the test sequence
CC adjacent to a defined protein-binding screening sequence. Binding
CC of moIs. to these test sequences changes the binding characteristics
CC of the protein mol. to its cognate binding sequence. When such a mol.
CC binds the test sequence, the equilibrium of the DNA:protein complexes
CC is disturbed, generating changes in the concentration of free DNA probe.
CC One application of this method is to eucaryotic general transcription
CC factors (e.g. TFIIID), where the target region is typically selected
CC from DNA sequences adjacent to the binding site for the eucaryotic
CC transcription factor. Numerous exemplary test sequences are given:
CC the sequences in 069251-731 and 069850 correspond to promoter targets
CC (typically TATA box-contg. sites) for human genes and the sequences in
CC 069732-849 correspond to promoter targets for viral genes. The test

CC sequences may also be randomly generated. DNA:protein interaction may
CC be used for screening purposes, e.g. the Herpes Simplex Virus (HSV)
CC origin of replication and U9 (see Q69851-52, Q69865 and Q69891).
SQ Sequence 47 BP; 13 A; 8 C; 10 G; 16 T;

Query Match 0.2%; Score 47; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12078 CAAATAGTGGTGGTCTACTCTATAAAGAAGAGTACTCTTTCCCTT 12124
1 CAAATAGTGGTGGTCTACTCTATAAAGAAGAGTACTCTTTCCCTT 47

RESULT 5
T63882 standard; DNA; 47 BP.
AC T63882;
DT 14-MAR-1997 (first entry)
DE Human H3 histone gene TFIID binding site.
KW Duplex DNA; target region; binding characteristic; DNA binding protein;
KW TFIID; transcription factor; binding site; inhibition; enhance; hCG;
KW cancer; inherited genetic disorder; alpha-D-galactosidase A; ds.
OS Homo sapiens.
PN US5578444-A.
PD 26-NOV-1996.
PF 27-JUN-1991; 723618.
PR 27-JUN-1991; US-723618.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PR 20-DEC-1993; US-171389.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
PT WPI: 97-020402/02.
PT Altering binding characteristics of DNA binding proteins to duplex
PT DNA - by attaching specific small cpd. to target region close to the
PT protein's binding site, useful in treatment of viral disease, cancer
PT etc
PS Claim 6; Column 185-186; 264pp; English.
CC The sequences given in T63713-4312 represent duplex DNA's which act
CC as target regions in the method of the invention. The method for
CC altering the binding characteristics of a DNA-binding protein to duplex
CC DNA comprises contacting the duplex DNA with a small molecule which
CC binds sequence-specifically to a target region, where, when the small
CC molecule is bound to the target region, it is adjacent to, but not
CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
CC The small molecule is added at a concentration effective to alter the
CC binding of the DNA binding protein, pref. TFIID, to its binding site on
CC the duplex DNA. The binding of the small molecule may inhibit or enhance
CC the binding of the DNA-binding protein to its binding site. The
CC compounds isolated using this method are potentially useful as
CC therapeutic agents for treatment of any disease which involves a
CC specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.
CC The method is suitable for screening large biological or chemical
CC libraries and allows determination of sequence-specific and relative
CC affinities of known DNA-binding agents for different DNA sequences.
CC The design of these duplex DNA's allows a single DNA:protein interaction
CC to be used for screening sequence-specific, or preferential, DNA binding
CC proteins that recognise almost any possible sequence (see also T49539-
CC 74).
SQ Sequence 47 BP; 13 A; 8 C; 10 G; 16 T;

Query Match 0.2%; Score 47; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12078 CAAATAGTGGTGGTCTACTCTATAAAGAAGAGTACTCTTTCCCTT 12124
1 CAAATAGTGGTGGTCTACTCTATAAAGAAGAGTACTCTTTCCCTT 47

RESULT 6
X17170 standard; DNA; 47 BP.
ID X17170;
AC X17170;
DT 06-MAY-1999 (first entry)
DE Test sequence from human histone H3 gene.
KW Test sequence; DNA-binding molecule; screening sequence; human;
KW nucleic acid amplification; target; viral; ds.
OS Homo sapiens.
PN US869241-A.
PD 09-FEB-1999.
PF 07-JUN-1995; 475228.
PR 20-DEC-1993; US-171389.
PR 27-JUN-1991; US-723618.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PR 07-JUN-1995; US-475228.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
PT WPI: 99-152755/13

PT Determination of DNA sequence preference of a DNA-binding molecule -
PT based on inhibition of binding of protein to oligonucleotide
PT sequence attached to test sequence
PS Claim 3; Columns 187-188; 270pp; English.
CC Sequences X17001 to X17600 represent specifically claimed target test
CC DNA sequence preference of a DNA-binding molecule. The method comprises:
CC (1) adding a test molecule and a DNA-binding protein to a mixture of
CC duplex DNA test oligonucleotides, each of the test oligonucleotides
CC having a test sequence adjacent to a screening sequence, where the
CC screening sequence binds to the DNA-binding protein with a binding
CC affinity that is independent of the DNA sequence of the test sequence,
CC and where the mixture of duplex DNA test oligonucleotides includes
CC several test sequences; (11) incubating the test molecule, the mixture of
CC duplex DNA test oligonucleotides and the DNA-binding protein for a time
CC sufficient to permit binding of the test molecule to test sequences in
CC the duplex DNA; (111) separating unbound test oligonucleotides from test
CC oligonucleotides bound to binding protein; (1v) amplifying the unbound
CC test oligonucleotides; (v) repeating steps (1) to (1v); (vi) isolating
CC the amplified test oligonucleotides; and (vii) sequencing the isolated
CC test oligonucleotides. Test sequences X17001-X17481 and X17600 correspond
CC to promoter targets for human genes and test sequences X17482-X17599
CC correspond to promoter targets for viral genes.
SQ Sequence 47 BP; 13 A; 8 C; 10 G; 16 T;

Query Match 0.2%; Score 47; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12078 CAAATAGTGGTGGTCTACTCTATAAAGAAGAGTACTCTTTCCCTT 12124
1 CAAATAGTGGTGGTCTACTCTATAAAGAAGAGTACTCTTTCCCTT 47

RESULT 7
Q69417 standard; DNA; 46 BP.
ID Q69417;
AC Q69417;
DT 27-FEB-1995 (first entry)
DE Human H2A histone gene, target region.
KW DNA protein-binding assay; test sequence; screening sequence;
KW promoter; target; TATA box; Herpes Simplex Virus; HSV;
KW origin of replication; U9; transcription factor; TFIID; ds.
OS Synthetic.
PN W09414980-A.
PD 07-JUL-1994.
PF 23-DEC-1992; US-996783.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
PT WPI: 94-234711/28.

PT Sequence-directed DNA-binding molecules - useful in
PT pharmaceuticals and as molecular reagents
PS Claim 28, Page 295, 587pp; English.
CC A DNA protein-binding assay is provided, useful for screening
CC libraries of synthetic or biological cpts. for their ability
CC to bind DNA test sequences. The assay is versatile in that any
CC number of test sequences can be tested by placing the test sequence
CC adjacent to a defined protein-binding screening sequence. Binding
CC of mols. to these test sequences changes the binding characteristics
CC of the protein mol. to its cognate binding sequence. When such a mol.
CC binds the test sequence, the equilibrium of the DNA:protein complexes
CC is disturbed, generating changes in the concentration of free DNA probe.
CC One application of this method is to eucaryotic general transcription
CC factors (e.g., TFIID), where the target region is typically selected
CC from DNA sequences adjacent to the binding site for the eucaryotic
CC transcription factor. Numerous exemplary test sequences are given:
CC the sequences in 069251-731 and 069850 correspond to promoter targets
CC (typically, TATA box-contg. sites) for human genes and the sequences in
CC 069732-849 correspond to promoter targets for viral genes. The test
CC sequences may also be randomly generated. DNA:protein interaction may
CC be used for screening purposes, e.g., the Herpes Simplex Virus (HSV)
CC origin of replication and US9 (see 069851-52, 069855 and 069891).
SQ Sequence 46 BP; 10 A; 16 C; 6 G; 14 T;

Query Match 0.2%; Score 46; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10562 CACAGCTTACTCTCAGTATTAATTAATCTCTGCTGCGTTC 10607
DB 1 CACAGCTTACTCTCAGTATTAATTAATCTCTGCTGCGTTC 46

RESULT 8
ID T63879 standard; DNA; 46 BP.
AC T63879.
DE Human H2A histone gene TFIID binding site.
KW Duplex DNA; target region; binding characteristic; DNA binding protein;
KW TFIID; transcription factor; binding site; inhibition; enhance; hcg;
KW cancer; inherited genetic disorder; alpha-D-galactosidase A; ds.
OS Homo sapiens.
PN US5578444-A.
PD 26-NOV-1996.
PR 27-JUN-1991; US-723618.
PR 27-JUN-1991; US-723618.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PR 20-DEC-1993; US-171389.
PI (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
PI WPI: 99-020402/02.
PT Altering binding characteristics of DNA binding proteins to duplex
PT DNA - by attaching specific small cpd. to target region close to the
PT protein's binding site, useful in treatment of viral disease, cancer
PT etc.
PS Claim 6; Column 183-184; 264pp; English.
CC The sequences given in T63713-4312 represent duplex DNA's which act
CC as target regions in the method of the invention. The method for
CC altering the binding characteristics of a DNA-binding protein to duplex
CC DNA comprises contacting the duplex DNA with a small molecule which
CC binds sequence-specifically to a target region, where, when the small
CC molecule is bound to the target region, it is adjacent to, but not
CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
CC The small molecule is added at a concentration effective to alter the
CC binding of the DNA binding protein, pref. TFIID, to its binding site on
CC the duplex DNA. The binding of the small molecule may inhibit or enhance
CC the binding of the DNA-binding protein to its binding site. The
CC compounds isolated using this method are potentially useful as
CC therapeutic agents for treatment of any disease which involves a
CC specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.

CC The method is suitable for screening large biological or chemical
CC libraries and allows determination of sequence-specific and relative
CC affinities of known DNA-binding agents for different DNA sequences.
CC The design of these duplex DNA's allows a single DNA:protein interaction
CC to be used for screening sequence-specific, or preferential, DNA binding
CC proteins that recognise almost any possible sequence (see also T63539-
CC 74).
SQ Sequence 46 BP; 10 A; 16 C; 6 G; 14 T;

Query Match 0.2%; Score 46; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10562 CACAGCTTACTCTCAGTATTAATTAATCTCTGCTGCGTTC 10607
DB 1 CACAGCTTACTCTCAGTATTAATTAATCTCTGCTGCGTTC 46

RESULT 9
ID X17167 standard; DNA; 46 BP.
AC X17167.
DE 06-MAY-1999 (first entry)
DE Test sequence from human histone H2A gene.
KW Test sequence; DNA-binding molecule; screening sequence; human;
KW nucleic acid amplification; target; viral; ds.
OS Homo sapiens.
PN US5869241-A.
PD 09-FEB-1999.
PR 07-JUN-1993; 475228.
PR 20-DEC-1993; US-171389.
PR 27-JUN-1991; US-723618.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PR 07-JUN-1995; US-475228.
PI (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
PI WPI: 99-152755/13.
PT Determination of DNA sequence preference of a DNA-binding molecule -
PT based on inhibition of binding of protein to oligonucleotide
PT sequence attached to test sequence
PS Claim 3; Columns 185-186; 270pp; English.
CC Sequences X17001 to X17600 represent specifically claimed target test
CC DNA sequences that are used in the method of the invention of determining the
CC DNA sequence preference of a DNA-binding molecule. The method comprises:
CC (1) adding a test molecule and a DNA-binding protein to a mixture of
CC duplex DNA test oligonucleotides, each of the test oligonucleotides
CC having a test sequence adjacent to a screening sequence, where the
CC screening sequence binds to the DNA-binding protein with a binding
CC affinity that is independent of the DNA sequence of the test sequence,
CC and where the mixture of duplex DNA test oligonucleotides includes
CC several test sequences; (11) incubating the test molecule, the mixture of
CC duplex DNA test oligonucleotides and the DNA-binding protein for a time
CC sufficient to permit binding of the test molecule to test sequences in
CC the duplex DNA; (111) separating unbound test oligonucleotides from test
CC oligonucleotides bound to binding protein; (11v) amplifying the unbound
CC test oligonucleotides; (v) repeating steps (11) to (11v); (v1) isolating
CC the amplified test oligonucleotides; and (v11) sequencing the isolated
CC test oligonucleotides. Test sequences X17001-X17481 and X17600 correspond
CC to promoter targets for human genes and test sequences X17482-X17599
CC correspond to promoter targets for viral genes.
SQ Sequence 46 BP; 10 A; 16 C; 6 G; 14 T;

Query Match 0.2%; Score 46; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10562 CACAGCTTACTCTCAGTATTAATTAATCTCTGCTGCGTTC 10607
DB 1 CACAGCTTACTCTCAGTATTAATTAATCTCTGCTGCGTTC 46

RESULT 10
ID 069419 standard; DNA: 45 BP.
AC 069419;
DE 27-FEB-1995 (first entry)
KW Human H3 histone gene, target region.
KW DNA protein-binding assay; test sequence; screening sequence;
KW promoter; target; TATA box; Herpes Simplex Virus; HSV;
KW origin of replication; UL9; transcription factor; TFIID; ds.
OS Synthetic.
PN WO9414980-A.
PD 07-JUL-1994.
PF 20-DEC-1993; U12388.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
DR WPI: 94-234711/28
PT Sequence-directed DNA-binding molecules - useful in
PT pharmaceuticals and as molecular reagents
PS Claim 28; Page 296; 587pp; English.
CC A DNA protein-binding assay is provided, useful for screening
CC libraries of synthetic or biological cpds. for their ability
CC to bind DNA test sequences. The assay is versatile in that any
CC number of test sequences can be tested by placing the test sequence
CC adjacent to a defined protein-binding screening sequence. Binding
CC of mols. to these test sequences changes the binding characteristics
CC of the protein mol. to its cognate binding sequence. When such a mol.
CC binds the test sequence, the equilibrium of the DNA:protein complexes
CC is disturbed, generating changes in the concentration of free DNA probe.
CC One application of this method is to eucaryotic general transcription
CC factors (e.g. TFIID), where the target region is typically selected
CC from DNA sequences adjacent to the binding site for the eucaryotic
CC transcription factor. Numerous exemplary test sequences are given:
CC the sequences in 069251-731 and 069850 correspond to promoter targets
CC (typically, TATA box-contg. sites) for human genes and the sequences in
CC 069732-849 correspond to promoter targets for viral genes. The test
CC sequences may also be randomly generated. DNA:protein interaction may
CC be used for screening purposes, e.g. the Herpes Simplex Virus (HSV)
CC origin of replication and UL9 (see 069851-52, 069865 and 069891).
SQ Sequence 45 BP; 8 A; 6 C; 10 G; 21 T;

Query Match 0.1%; Score 45; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12057 ATTTTGAATTTCTGGGTCATAGTGGTGTGACTCTAT 12101
DB 1 ATTTTGAATTTCTGGGTCATAGTGGTGTGACTCTAT 45

RESULT 11
ID T63881 standard; DNA: 45 BP.
AC T63881;
DE 14-MAR-1997 (first entry)
KW Human H3 histone gene TFIID binding site.
KW Duplex DNA; target region; binding characteristic; DNA binding protein;
KW TFIID; transcription factor; binding site; inhibition; enhancer; hCG;
KW cancer; inherited genetic disorder; alpha-D-galactosidase A; ds.
OS Homo sapiens.
PN US5578444-A.
PD 26-NOV-1996.
PF 27-JUN-1991; 723618.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PR 20-DEC-1993; US-171389.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
DR WPI: 97-020402/02.

PT Altering binding characteristics of DNA binding proteins to duplex
PT DNA - by attaching specific small cpd. to target region close to the
PT protein's binding site, useful in treatment of viral disease, cancer
PT etc
PS Claim 6; Column 185-186; 264pp; English.
CC The sequences given in T63713-4312 represent duplex DNA's which act
CC as target regions in the method of the invention. The method for
CC altering the binding characteristics of a DNA-binding protein to duplex
CC DNA comprises contacting the duplex DNA with a small molecule which
CC binds sequence-specifically to a target region, where, when the small
CC molecule is bound to the target region, it is adjacent to, but not
CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
CC The small molecule is added at a concentration effective to alter the
CC binding of the DNA binding protein, pref. TFIID, to its binding site on
CC the duplex DNA. The binding of the small molecule may inhibit or enhance
CC the binding of the DNA-binding protein to its binding site. The
CC compounds isolated using this method are potentially useful as
CC therapeutic agents for treatment of any disease which involves a
CC specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.
CC The method is suitable for screening large biological or chemical
CC libraries and allows determination of sequence-specific and relative
CC affinities of known DNA-binding agents for different DNA sequences.
CC The design of these duplex DNA's allows a single DNA:protein interaction
CC to be used for screening sequence-specific, or preferential, DNA binding
CC proteins that recognise almost any possible sequence (see also T49539-
CC 74).
SQ Sequence 45 BP; 8 A; 6 C; 10 G; 21 T;

Query Match 0.1%; Score 45; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12057 ATTTTGAATTTCTGGGTCATAGTGGTGTGACTCTAT 12101
DB 1 ATTTTGAATTTCTGGGTCATAGTGGTGTGACTCTAT 45

RESULT 12
ID X17169 standard; DNA: 45 BP.
AC X17169;
DE 06-MAY-1999 (first entry)
KW Test sequence; DNA-binding molecule; screening sequence; human;
KW nucleic acid amplification; target; viral; ds.
OS Homo sapiens.
PN US5869241-A.
PD 09-FEB-1999.
PF 07-JUN-1995; 475228.
PR 20-DEC-1993; US-171389.
PR 27-JUN-1991; US-723618.
PR 23-DEC-1992; US-996783.
PR 17-SEP-1993; US-123936.
PR 07-JUN-1995; US-475228.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
DR WPI: 99-152735/13.
PT Determination of DNA sequence preference of a DNA-binding molecule -
PT based on inhibition of binding of protein to oligonucleotide
PT sequence attached to test sequence
PS Claim 3; Columns 185-186; 270pp; English.
CC Sequences X17001 to X17600 represent specifically claimed target test
CC sequences that are used in the method of the invention of determining the
CC DNA sequence preference of a DNA-binding molecule. The method comprises:
CC (1) adding a test molecule and a DNA-binding protein to a mixture of
CC duplex DNA test oligonucleotides, each of the test oligonucleotides
CC having a test sequence adjacent to a screening sequence, where the
CC screening sequence binds to the DNA-binding protein with a binding
CC affinity that is independent of the DNA sequence of the test sequence,
CC and where the mixture of duplex DNA test oligonucleotides includes
CC several test sequences; (11) incubating the test molecule, the mixture of
CC duplex DNA test oligonucleotides and the DNA-binding protein for a time

CC sufficient to permit binding of the test molecule to test sequences in
CC the duplex DNA; (iii) separating unbound test oligonucleotides from test
CC oligonucleotides bound to binding protein; (iv) amplifying the unbound
CC test oligonucleotides; (v) repeating steps (ii) to (iv); (vi) isolating
CC the amplified test oligonucleotides; and (vii) sequencing the isolated
CC test oligonucleotides. Test sequences X17001-X17481 and X17600 correspond
CC to promoter targets for human genes and test sequences X17482-X17539
CC correspond to promoter targets for viral genes
SQ Sequence 45 BP; 8 A; 6 C; 10 G; 21 T;

Query Match 0.1%; Score 45; DB 1; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12057 ATTTTGAATTTCTTGGTCCAAATAGTGTGCTGACTCTAT 12101
DB 1 ATTTTGAATTTCTTGGTCCAAATAGTGTGCTGACTCTAT 45

RESULT 13
062216/c
ID 062216 standard; DNA; 30 BP.
AC 062216;
DT 22-NOV-1994 (first entry)
DE Histone H3 mRNA probe #1.
KW Probe: histone; H2D; H3; H4; mRNA; detection; proliferation; cell;
OS Synthesis; cancer; carcinogen; ss.
PN WO9408046-A.
PD 14-APR-1994.
PF 20-SEP-1993; GB-020777.
PR 02-OCT-1992; UNIV LEICESTER.
PA (UYLE-) UNIV LEICESTER.
PI Pringle JH;
PI WPI: 94-135606/16.
PT Detecting proliferating cells in tissue or cell samples - by
PT hybridisation with non-radioactively labelled oligonucleotide
PT probes specific for histone mRNA.
PS Claim 13; Page 18; 24pp; English.
CC The sequences given in 062216-18 are probes which are specific for
CC histone H3 mRNA. These probes may be used in the method of the
CC invention for detecting the presence of proliferating cells in a
CC tissue/cell sample. The method comprises preparing the tissue/cell
CC sample, hybridising these non-radioactively labelled probes, and
CC detecting the presence of histone-probe hybrids. The method can be
CC used to identify cells undergoing proliferation in normal and
CC diseased tissue. It can be used in the diagnosis of cancers and in
CC assessing the carcinogenic properties of a chemical.
SQ Sequence 30 BP; 6 A; 6 C; 8 G; 10 T;

Query Match 0.1%; Score 30; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12518 CGAGTGACTATATGCCCCAAGACATCCAG 12547
DB 30 CGAGTGACTATATGCCCCAAGACATCCAG 1

RESULT 14
062217/c
ID 062217 standard; DNA; 30 BP.
AC 062217;
DT 22-NOV-1994 (first entry)
DE Histone H3 mRNA probe #2.
KW Probe: histone; H2D; H3; H4; mRNA; detection; proliferation; cell;
OS Synthesis; cancer; carcinogen; ss.
PN WO9408046-A.
PD 14-APR-1994.
PF 20-SEP-1993; GB-01977.

PR 02-OCT-1992; GB-020777.
PA (UYLE-) UNIV LEICESTER.
PI Pringle JH;
PI WPI: 94-135606/16.
PT Detecting proliferating cells in tissue or cell samples - by
PT hybridisation with non-radioactively labelled oligonucleotide
PT probes specific for histone mRNA.
PS Claim 13; Page 18; 24pp; English.
CC The sequences given in 062216-18 are probes which are specific for
CC histone H3 mRNA. These probes may be used in the method of the
CC invention for detecting the presence of proliferating cells in a
CC tissue/cell sample. The method comprises preparing the tissue/cell
CC sample, hybridising these non-radioactively labelled probes, and
CC detecting the presence of histone-probe hybrids. The method can be
CC used to identify cells undergoing proliferation in normal and
CC diseased tissue. It can be used in the diagnosis of cancers and in
CC assessing the carcinogenic properties of a chemical.
SQ Sequence 30 BP; 7 A; 7 C; 8 G; 8 T;

Query Match 0.1%; Score 30; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12330 GCTACCAAAAGTCGACGAGTTCGTATTC 12359
DB 30 GCTACCAAAAGTCGACGAGTTCGTATTC 1

RESULT 15
062218/c
ID 062218 standard; DNA; 30 BP.
AC 062218;
DT 22-NOV-1994 (first entry)
DE Histone H3 mRNA probe #3.
KW Probe: histone; H2D; H3; H4; mRNA; detection; proliferation; cell;
OS Synthesis; cancer; carcinogen; ss.
PN WO9408046-A.
PD 14-APR-1994.
PF 20-SEP-1993; GB-020777.
PR 02-OCT-1992; UNIV LEICESTER.
PA (UYLE-) UNIV LEICESTER.
PI Pringle JH;
PI WPI: 94-135606/16.
PT Detecting proliferating cells in tissue or cell samples - by
PT hybridisation with non-radioactively labelled oligonucleotide
PT probes specific for histone mRNA.
PS Claim 13; Page 18; 24pp; English.
CC The sequences given in 062216-18 are probes which are specific for
CC histone H3 mRNA. These probes may be used in the method of the
CC invention for detecting the presence of proliferating cells in a
CC tissue/cell sample. The method comprises preparing the tissue/cell
CC sample, hybridising these non-radioactively labelled probes, and
CC detecting the presence of histone-probe hybrids. The method can be
CC used to identify cells undergoing proliferation in normal and
CC diseased tissue. It can be used in the diagnosis of cancers and in
CC assessing the carcinogenic properties of a chemical.
SQ Sequence 30 BP; 4 A; 6 C; 9 G; 11 T;

Query Match 0.1%; Score 30; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12177 GTACTAAACAGACAGCTGGAAATCCACCG 12206
DB 30 GTACTAAACAGACAGCTGGAAATCCACCG 1

Search completed: March 30, 2000, 17:38:54
Job time: 136132 sec

Mon Apr 3 08:24:32 2000

us-08-852-495c-1_copy_90000_120000.png

Page 7

Mon Apr 3 08:24:33 2000

us-08-852-495c-1_copy_90000_120000.rn1

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 10:42:50 ; Search time 344.26 Seconds
(without alignments)
10433.435 Million cell updates/sec

Title: US-08-852-495C-1_COPY_90000_120000

Perfect score: 30001
Sequence: 1 TGGCGAGCTGAGTGCGCG.....TATCTAGCATGTCAT 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 214294 seqs, 59861574 residues

Word size : 0

Total number of hits satisfying chosen parameters: 349160

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database :

- 1: Issued Patents.NA.*
- 2: /cgn2_6/prodata/1/lna/5A_COMB.seq.*
- 3: /cgn2_6/prodata/1/lna/5C_COMB.seq.*
- 4: /cgn2_6/prodata/1/lna/5D_COMB.seq.*
- 5: /cgn2_6/prodata/1/lna/6_COMB.seq.*
- 6: /cgn2_6/prodata/1/lna/PCTUS9_COMB.seq.*
- 7: /cgn2_6/prodata/1/lna/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58	0.2	76	3	US-08-454-557C-69
2	58	0.2	76	4	US-08-340-426D-69
3	58	0.2	76	4	US-08-450-673C-69
4	58	0.2	76	6	PCT-US95-17111A-69
5	55	0.2	60	3	US-08-454-557C-57
6	55	0.2	60	4	US-08-340-426D-57
7	55	0.2	60	4	US-08-450-673C-57
8	55	0.2	60	6	PCT-US95-17111A-57
9	48	0.2	48	1	US-08-171-389-159
10	48	0.2	48	2	US-08-123-936-159
11	48	0.2	48	3	US-08-475-228A-159
12	48	0.2	48	5	US-08-482-080A-159
13	48	0.2	48	6	PCT-US93-12388-159
14	47	0.2	47	1	US-08-171-389-170
15	47	0.2	47	2	US-08-123-936-170
16	47	0.2	47	3	US-08-475-228A-170
17	47	0.2	47	5	US-08-482-080A-170
18	47	0.2	47	6	PCT-US93-12388-170
19	46	0.2	46	1	US-08-171-389-167
20	46	0.2	46	2	US-08-123-936-167
21	46	0.2	46	3	US-08-475-228A-167
22	46	0.2	46	5	US-08-482-080A-167
23	46	0.2	46	6	PCT-US93-12388-167
24	45	0.1	45	1	US-08-171-389-169
25	45	0.1	45	2	US-08-123-936-169
26	45	0.1	45	3	US-08-475-228A-169

27	45	0.1	45	5	US-08-482-080A-169	Sequence 169, App
28	45	0.1	45	6	PCT-US93-12388-169	Sequence 169, App
29	28	0.1	30	4	US-08-859-998-318	Sequence 318, App
30	28	0.1	66	3	US-08-454-557C-67	Sequence 67, App1
31	28	0.1	66	4	US-08-340-426D-67	Sequence 67, App1
32	28	0.1	66	4	US-08-450-673C-67	Sequence 67, App1
33	28	0.1	66	6	PCT-US95-17111A-67	Sequence 67, App1
34	28	0.1	85	3	US-08-454-557C-92	Sequence 92, App1
35	28	0.1	85	4	US-08-340-426D-92	Sequence 92, App1
36	28	0.1	85	4	US-08-450-673C-92	Sequence 92, App1
37	28	0.1	85	6	PCT-US95-17111A-92	Sequence 92, App1
38	27	0.1	36	3	US-08-669-721-5	Sequence 5, App1
39	27	0.1	57	1	US-08-222-177A-163	Sequence 163, App1
40	27	0.1	84	3	US-08-454-557C-91	Sequence 91, App1
41	27	0.1	84	4	US-08-340-426D-91	Sequence 91, App1
42	27	0.1	84	4	US-08-450-673C-91	Sequence 91, App1
43	27	0.1	84	6	PCT-US95-17111A-91	Sequence 91, App1
44	26	0.1	35	1	US-08-255-889-10	Sequence 10, App1
45	26	0.1	36	3	US-08-669-721-5	Sequence 5, App1

ALIGNMENTS

RESULT 1
US-08-454-557C-69
Sequence 69, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-69
Query Match 0.2% Score 58; DB 3; Length 76;
Best Local Similarity 100.0% Pred. No. 2.1e-10;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2685 GCCCGCTAATTTTATTTAGTAGACAGGCTTCACCGTGTGGCCAGATG 2742
DB 6 GCCCGCTAATTTTATTTAGTAGACAGGCTTCACCGTGTGGCCAGATG 63

Mon Apr 3 08:24:33 2000

us-08-852-495c-1_copy_90000_120000.rml

Page 2

```
RESULT 2
US-08-340-426D-69
; Sequence 69, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-340-426D-69

Query Match          0.28; Score 58; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2685 GCCCGCGTAATTTTGTATTTTGTATTTGATGAGACAGGAGGTTTCCACCGTGTGGCCAGAGT 2742
DB 6 GCCCGCGTAATTTTGTATTTTGTATTTGATGAGACAGGAGGTTTCCACCGTGTGGCCAGAGT 63

RESULT 3
US-08-450-673C-69
; Sequence 69, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-450-673C-69
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-69

Query Match          0.28; Score 58; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2685 GCCCGCGTAATTTTGTATTTTGTATTTGATGAGACAGGAGGTTTCCACCGTGTGGCCAGAGT 2742
DB 6 GCCCGCGTAATTTTGTATTTTGTATTTGATGAGACAGGAGGTTTCCACCGTGTGGCCAGAGT 63

RESULT 4
PCT-US95-17111A-69
; Sequence 69, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
PCT-US95-17111A-69
```

Query Match 0.2%; Score 58; DB 6; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2685 GCCCGCTAATTTTGTATTTAGTAGAGACAGGGTTCCACCGTGTGGCCAGGATG 2742
DB 6 GCCCGCTAATTTTGTATTTAGTAGAGACAGGGTTCCACCGTGTGGCCAGGATG 63

RESULT 5

US-08-454-557C-57
; Sequence 57, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-454-557C-57

Query Match 0.2%; Score 55; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2685 GCCCGCTAATTTTGTATTTAGTAGAGACAGGGTTCCACCGTGTGGCCAGG 2739
DB 6 GCCCGCTAATTTTGTATTTAGTAGAGACAGGGTTCCACCGTGTGGCCAGG 60

RESULT 6

US-08-340-426D-57
; Sequence 57, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-57

Query Match 0.2%; Score 55; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2685 GCCCGCTAATTTTGTATTTAGTAGAGACAGGGTTCCACCGTGTGGCCAGG 2739
DB 6 GCCCGCTAATTTTGTATTTAGTAGAGACAGGGTTCCACCGTGTGGCCAGG 60

RESULT 7

US-08-450-673C-57
; Sequence 57, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 57;
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-57

Query Match 0.2%; Score 55; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2685 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGTTTCACCGTGTGGCCAGG 2739
DB 6 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGTTTCACCGTGTGGCCAGG 60

RESULT 8
PCT-US95-1711A-57
Sequence 57, Application PC/TUS951711A
GENERAL INFORMATION:
APPLICANT: de la Monte, Susanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/1711A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-1711A-57

Query Match 0.2%; Score 55; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 2e-09;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2685 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGTTTCACCGTGTGGCCAGG 2739
DB 6 GCCCGCGTAATTTTGTATTTAGTAGAGACAGGTTTCACCGTGTGGCCAGG 60

RESULT 9
US-08-171-389-159/C

Sequence 159, Application US/08171389
Patent No. 5578444
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fry, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
NUMBER OF SEQUENCES: 641
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/171,389
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0175/G19P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 159:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human H4/a gene for H4 histone
US-08-171-389-159

Query Match 0.2%; Score 48; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22582 GAGAGATAGGAACCGATCTTATATCTAGTACCGCCGCCGCC 22629
DB 48 GAGAGATAGGAACCGATCTTATATCTAGTACCGCCGCCGCC 1

RESULT 10
US-08-123-936-159/C
Sequence 159, Application US/08123936
Patent No. 5726014
GENERAL INFORMATION:

APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
TITLE OF INVENTION: Screening Assay for the Detection of
TITLE OF INVENTION: DNA-Binding Molecules
NUMBER OF SEQUENCES: 640
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/123,936
FILING DATE: 27-JUN-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Fadian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 159:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human H4/a gene for H4 histone
US-08-123-936-159

Query Match 0.2%; Score 48; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22582 GAGAGATAGGAACGATCTTTATATCTACGTTACCCCTGCCCC 22629
Db 48 GAGAGATAGGAACGATCTTTATATCTACGTTACCCCTGCCCC 1

RESULT 11
US-08-475-228A-159/c
Sequence 159, Application US/08475228A
Patent No. 5869241
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fry, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 664
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.

STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,228A
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Stratford, Carol A.
REGISTRATION NUMBER: 34,444
REFERENCE/DOCKET NUMBER: 4600-0175.21/G19P3D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 159:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human H4/a gene for H4 histone
US-08-475-228A-159

Query Match 0.2%; Score 48; DB 3; Length 48;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22582 GAGAGATAGGAACGATCTTTATATCTACGTTACCCCTGCCCC 22629
Db 48 GAGAGATAGGAACGATCTTTATATCTACGTTACCCCTGCCCC 1

RESULT 12
US-08-482-080A-159/c
Sequence 159, Application US/08482080A
Patent No. 6010849
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fry, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 664
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA

Mon Apr 3 08:24:33 2000

us-08-852-495c-1_copy_90000_120000.inl

Page 6

```

? ZIP: 94063
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? FILING DATE: 07-JUN-1993
? APPLICATION NUMBER: US 08/483,080A
? PRIORITY APPLICATION DATA:
? FILING DATE: 20-DEC-1993
? APPLICATION NUMBER: US 08/171,389
? PRIORITY APPLICATION DATA:
? FILING DATE: 17-SEP-1993
? APPLICATION NUMBER: US 08/123,936
? PRIORITY APPLICATION DATA:
? FILING DATE: 23-DEC-1992
? APPLICATION NUMBER: US 07/996,783
? PRIORITY APPLICATION DATA:
? FILING DATE: 27-JUN-1991
? APPLICATION NUMBER: US 07/723,618
? PRIORITY APPLICATION DATA:
? FILING DATE: 22-JUN-1993
? APPLICATION NUMBER: US 08/081,070
? ATTORNEY/AGENT INFORMATION:
? NAME: Brady, John F.
? REGISTRATION NUMBER: 39,118
? REFERENCE/DOCKET NUMBER: 4600-0175.20/G19P3D1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (650) 324-0980
? TELEFAX: (650) 324-0960
? INFORMATION FOR SEQ ID NO: 159:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 48 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? HYPOTHEICAL: NO
? ORIGINAL SOURCE:
? INDIVIDUAL ISOLATE: Human H4/a gene for H4 histone
? US-08-482-080A-159

Query Match 0.28; Score 48; DB 5; Length 48;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22582 GAGAGAAATGGAACCGATCTTTATATCTACGTTACCCCTGCCCCC 22629
Db 48 GAGAGAAATGGAACCGATCTTTATATCTACGTTACCCCTGCCCCC 1

RESULT 13
PCT-US93-12388-159/C
? Sequence 159, Application PC/TUS9312388
? GENERAL INFORMATION:
? APPLICANT:
? TITLE OF INVENTION: Sequence-directed DNA Binding
? NUMBER OF SEQUENCES: 641
? CORRESPONDENCE ADDRESSES:
? ADDRESSEE: Genelabs Technologies, Inc.
? STREET: 505 Penobscot Drive
? CITY: Redwood City
? STATE: CA
? COUNTRY: USA
? ZIP: 94063
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25

```

```

? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: PCT/US93/12388
? FILING DATE:
? CLASSIFICATION:
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 08/123,936
? FILING DATE: 17-SEP-1993
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 07/996,783
? FILING DATE: 23-DEC-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: Fabian Gatt, R.
? REGISTRATION NUMBER: 33,875
? REFERENCE/DOCKET NUMBER: 4600-0175.41/G19PCT2
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 324-0880
? TELEFAX: (415) 324-0960
? INFORMATION FOR SEQ ID NO: 159:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 48 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? HYPOTHEICAL: NO
? ORIGINAL SOURCE:
? INDIVIDUAL ISOLATE: Human H4/a gene for H4 histone
? PCT-US93-12388-159

Query Match 0.28; Score 48; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22582 GAGAGAAATGGAACCGATCTTTATATCTACGTTACCCCTGCCCCC 22629
Db 48 GAGAGAAATGGAACCGATCTTTATATCTACGTTACCCCTGCCCCC 1

RESULT 14
US-08-171-389-170
? Sequence 170, Application US/08171389
? Patent No. 5578444
? GENERAL INFORMATION:
? APPLICANT: Edwards, Cynthia A.
? APPLICANT: Cantor, Charles R.
? APPLICANT: Andrews, Beth M.
? APPLICANT: Turin, Lisa M.
? APPLICANT: Fry, Kirk E.
? TITLE OF INVENTION: Sequence-directed DNA Binding
? NUMBER OF SEQUENCES: 641
? CORRESPONDENCE ADDRESSES:
? ADDRESSEE: Genelabs Technologies, Inc.
? STREET: 505 Penobscot Drive
? CITY: Redwood City
? STATE: CA
? COUNTRY: USA
? ZIP: 94063
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/171,389
? FILING DATE:
? CLASSIFICATION: 435
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 08/123,936
? FILING DATE: 17-SEP-1993
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: US 07/996,783

```

FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0175/G19P3
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human histone H3 gene
US-08-171-389-170

Query Match 0.2%; Score 47; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12078 CAAATAGTGGTGTGACTCTATAAAGAGAGTAGCTCTTCCCT 12124
DB 1 CAAATAGTGGTGTGACTCTATAAAGAGAGTAGCTCTTCCCT 47

RESULT 15
US-08-123-936-170
Sequence 170, Application US/08123936
Patent No. 5726014
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
TITLE OF INVENTION: Screening Assay for the Detection of
NUMBER OF SEQUENCES: 640
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/123,936
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human histone H3 gene
US-08-123-936-170

Query Match 0.2%; Score 47; DB 2; Length 47;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12078 CAAATAGTGGTGTGACTCTATAAAGAGAGTAGCTCTTCCCT 12124
DB 1 CAAATAGTGGTGTGACTCTATAAAGAGAGTAGCTCTTCCCT 47

Search completed: March 30, 2000, 17:28:35
Job time: 135577 sec

Mon Apr 3 08:24:33 2000

us-08-852-495c-1_copy_90000_120000.rn1

Page 8

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 03:49:55 ; Search time 8198.5 Seconds

(without alignments)
13816.393 Million cell updates/sec

Title: US-08-852-495C-1_COPY_90000_120000

Perfect score: 30001
Sequence: 1 TGCCAGCTCAGTCAGTGGCG.....TATCTAAGCCATGTCAT 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : ESR:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*

44: gb_est25:*
45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: gb_est33:*
53: gb_est34:*
54: gb_est35:*
55: gb_est36:*
56: gb_est37:*
57: gb_est38:*
58: gb_est39:*
59: gb_est40:*
60: gb_est41:*
61: gb_est42:*
62: gb_est43:*
63: gb_est44:*
64: gb_est45:*
65: gb_est46:*
66: gb_est47:*
67: gb_est48:*
68: gb_est49:*
69: gb_est50:*
70: gb_est51:*
71: gb_est52:*
72: gb_est53:*
73: gb_est54:*
74: gb_est55:*
75: gb_est56:*
76: gb_est57:*
77: gb_est58:*
78: gb_est59:*
79: gb_est60:*
80: gb_est61:*
81: gb_est62:*
82: gb_est63:*
83: gb_est64:*
84: gb_est65:*
85: gb_est66:*
86: gb_est67:*
87: gb_est68:*
88: gb_est69:*
89: gb_est70:*
90: gb_est71:*
91: gb_est72:*
92: gb_est73:*
93: gb_est74:*
94: gb_est75:*
95: gb_est76:*
96: gb_est77:*
97: gb_est78:*
98: gb_est79:*
99: gb_est80:*
100: gb_est81:*
101: gb_est82:*
102: gb_est83:*
103: gb_est84:*
104: gb_est85:*
105: gb_est86:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
C 1	53	0.2	98 31 AA280198	AA280198 zt04b12.r

```
C 3 48 0.2 70 21 T64089
C 4 45 0.1 102 100 A0319270
C 5 43 0.1 91 50 F29274
C 6 42 0.1 95 34 A4457423
C 7 41 0.1 92 36 A1932967
C 8 40 0.1 104 105 A0626158
C 9 40 0.1 52 20 D25879
C 10 40 0.1 103 81 B48914
C 11 39 0.1 90 61 A1864984
C 12 35 0.1 35 60 A1801185
C 13 35 0.1 88 36 A809831
C 14 35 0.1 96 22 R67088
C 15 35 0.1 102 104 A0535244
C 16 35 0.1 103 104 A0582186
C 17 35 0.1 105 61 A1832832
C 18 34 0.1 57 36 A6534736
C 19 34 0.1 100 35 A578614
C 20 34 0.1 104 36 A614379
C 21 33 0.1 68 28 D44637
C 22 33 0.1 97 25 N49638
C 23 33 0.1 102 81 B48088
C 24 33 0.1 105 28 A084476
C 25 33 0.1 105 30 A247810
C 26 32 0.1 61 50 A1671034
C 27 32 0.1 95 44 A1281558
C 28 32 0.1 65 44 A828120
C 29 32 0.1 95 37 A669571
C 30 32 0.1 103 36 A807640
C 31 32 0.1 104 29 A128957
C 32 32 0.1 72 22 R70733
C 33 31 0.1 37 22 A4551166
C 34 31 0.1 71 35 A649287
C 35 31 0.1 82 36 A649287
C 36 31 0.1 90 20 Z25807
C 37 31 0.1 92 40 A916300
C 38 31 0.1 96 50 A1708433
C 39 31 0.1 98 24 H67549
C 40 31 0.1 101 30 A235077
C 41 31 0.1 101 39 A835205
C 42 31 0.1 102 36 A654562
C 43 30 0.1 34 25 N77004
C 44 30 0.1 60 37 A704865
C 45 30 0.1 81 B36140
```

ALIGNMENTS

```
RESULT 1
AA280198/c 98 bp mRNA EST 14-AUG-1997
LOCUS z004012.r1 NCI-CGAP-GCBI Homo sapiens cDNA clone IMAGE:712127 5'
DEFINITION similar to contatans Alu repetitive element; contains element MER22
repetitive element ; mRNA sequence.
ACCESSION AA280198
VERSION AA280198.1 GI:1921755
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 98)
NCT-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL On Sep 12 1996 this sequence version replaced gi:1395022.
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
```

```
FEATURES
Source Putative full length read
the vector to vector length is 99
Insert length: 622 Std Error: 0.00
Seq primer: -28m13 rev2Er from Amersham.
Location/Qualifiers
1..98
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:712127"
/clone_lib="NCI-CGAP-GCBI"
/tissue_type="germlinal center B cell"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germlinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI). Dr. David Allman
(NCI) and Dr. Gerald Marti (CEBR). cDNA synthesis was
primed with a Not I - oligo(dT) primer
[5'-TGTACCAATCTGAGGAGGAGCGCCGCTCATTTTCTTTTCTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
```

```
BASE COUNT 25 a 25 c 33 g 15 t
```

```
Query Match 0.2%; Score 53; DB 31; Length 98;
Best Local Similarity 100.0%; Pred. No. 7.3e-08;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2568 CCCAGCTGAGTGCATGCGCGATCTGCGTCACTGCACCTCCGCTTCCC 2620
DB 79 CCCAGCTGAGTGCATGCGCGATCTGCGTCACTGCACCTCCGCTTCCC 27
```

```
RESULT 2
T64089/c 70 bp mRNA EST 17-FEB-1995
LOCUS y005f08.r1 Stratiogene lung (#937210) Homo sapiens cDNA clone
DEFINITION IMAGE:7815 5' similar to contatans Alu repetitive element; mRNA
sequence.
ACCESSION T64089
VERSION T64089.1 GI:667954
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 70)
Hillier, L., Lennon, G., Becker, M., Donald, M.F., Chapell, B.,
Chissole, S., Dietrich, N., Dubuque, T., Favell, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, B.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.
and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence steps: 54
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: M13RP1
```

```
TITLE
JOURNAL
MEDLINE
COMMENT
```

FEATURES High quality sequence stop: 54.
Location/Qualifiers

source

1. .70

/organism="Homo sapiens"

/db_xref="GDB:483432"

/db_xref="taxon:9606"

/clone="IMAGE:79815"

/clone_1lb="Stratagene lung (#937210)"

/sex="male"

/dev_stage="72 years"

/lab_host="SOLR cells (Kanamycin resistant)"

/note="Organ: lung; Vector: pBluescript SK-; Site:1:
EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. normal lung. Average insert size: 1.0 kb;
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGCAG
3' -3' adaptor sequence: 5' CTCGATTTTCTTTTCTTTT 3' "

BASE COUNT

17 a 19 c 20 g 14 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 4,6e-06;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8371 TGCCAGCGGAGTGCAGTGCAGATCTCGGCTCAGTCGACCTCT 8418

Db 48 TGCCAGCGGAGTGCAGTGCAGATCTCGGCTCAGTCGACCTCT 1

RESULT 3

LOCUS

AP019270 102 bp DNA GSS 06-MAY-1999

DEFINITION RPII11-98B22.TV RPII-11 Homo sapiens genomic clone RPII-11-98B22,

genomic survey sequence.

ACCESSION

AP019270

VERSION

AP019270.1 GI:4052235

KEYWORDS

GSS.

SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 102)

Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,

Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.

Use of human BAC end sequences for Sequence-Ready Map Building

Unpublished (1998)

Other-GSS: RPII11-98B22.TV

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbeetfgr.org

Clones are derived from the human BAC library RPII-11. For BAC

library availability, please contact Pieter de Jong

(pieterdejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from

Research Genetics (info@resgen.com). BAC end search page:

http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html

Seg primer: SP6

Class: BAC ends.

Location/Qualifiers

source

1. .102

/organism="Homo sapiens"

/db_xref="GDB:7537293"

/db_xref="taxon:9606"

/clone="RPII-11-98B22"

/clone_1lb="RPII-11"

/sex="Male"

/cell_type="Lymphocytes"

/note="Vector: pBAC3.6; Site:1: EcoRI; Site:2: EcoRI;

RPII11 Human Male BAC Library"

BASE COUNT 26 a 28 c 31 g 17 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 4,1e-05;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3844 GGTGATCCACCCGCTCGGCTCCCAAGTCTGGATTACAGT 3888

Db 97 GGTGATCCACCCGCTCGGCTCCCAAGTCTGGATTACAGT 53

RESULT 4

LOCUS

F29274 F29274 91 bp mRNA EST 13-MAY-1999

DEFINITION HSPDI9063 HM3 Homo sapiens cDNA clone s4000074G04, mRNA sequence.

ACCESSION

F29274

VERSION

F29274.1 GI:4814900

KEYWORDS

EST.

SOURCE

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 91)

Lanfanchi,G., Muraro,T., Caldara,F., Pacchioni,B., Pallavicini,A.,

Pandolfi,D., Toppo,S., Trevisan,S., Scarso,S. and Valle,G.

Identification of 4370 expressed sequence tags from a

3'-end-specific cDNA library of human skeletal muscle by DNA

sequencing and filter hybridization

Genome Res. 6 (1), 35-42 (1996)

96276048

JOURNAL

MEDLINE

COMMENT

On Apr 7, 1998 this sequence version replaced gi:3034419.

Contact: Valle G.

CIRI Biotechnology Centre

University of Padua

Via Trieste 75, 35121 Padua, Italy

ABI Chromatograms and other information are available on WWW at

http://grup.bio.unipd.it.

Location/Qualifiers

source

1. .91

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="s4000074G04"

/clone_1lb="HM3"

/sex="female"

/tissue_type="pectoral muscle (after mastectomy)"

/note="Vector: pCDNAII (Invitrogen); Site:1: BstXI;

Site:2: NotI; The library was constructed by G.

Lanfanchi. This library is not substracted nor normalized.

The first strand cDNA was primed with a biotinylated

Oligo-dT-NotI primer

(5'-biotin-AACCGCTCGAGCGCGCTTTTCTTTTCTTTT-3'). The

ds cDNA was sonicated and size-selected in the range

350-550 bp. The 3' specific fragments were selected by

streptavidin coated magnetic beads, ligated to

non-palindromic BstXI adapters, NotI digested and

directionally cloned into BstXI-NotI cut pCDNAII vector."

BASE COUNT 18 a 19 c 38 g 16 t

ORIGIN

Query Match 0.1%; Score 43; DB 50; Length 91;

Best Local Similarity 100.0%; Pred. No. 0.00021;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1283 GGGCGCTGTAGTCCACGACTCGGAGGCTGAGGCGAGGAA 1325

Db 24 GGGCGCTGTAGTCCACGACTCGGAGGCTGAGGCGAGGAA 66

RESULT 5

AA457423/c

LOCUS AAA57423 95 bp mRNA EST 06-JUN-1997
 DEFINITION aa86002.r1 StrataGene fetal retina 937202 Homo sapiens cDNA clone
 IMAGE:838155 5' similar to contains Alu repetitive element; mRNA
 sequence.
 ACCESSION RA457423
 VERSION AA57423.1 GI:2180143
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 95)
 Haller, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
 Moore, B., Schellenberg, R., Steptoe, M., Tan, F., Theising, B.,
 White, Y., Wylie, T., Waterston, R. and Wilson, R.
 WashU-Merck EST Project 1997
 Unpublished (1997)
 On Nov 4, 1996 this sequence version replaced gi:1671271.
 TITLE JOURNAL
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -28ml3 rev1.1r from Amersham
 High quality sequence stop: 76.
 Location/Qualifiers
 1..95
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:838155"
 /clone_lib="Stratagene fetal retina 937202"
 /sex="mixed"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled
 retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
 Vector: -5' adaptor sequence: 5' GAATTCGCGCAGG 3' -3'
 adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'"
 BASE COUNT 14 a 39 c 25 g 17 t
 ORIGIN
 Query Match 0.1%; Score 42; DB 34; Length 95;
 Best Local Similarity 100.0%; Pred. No. 0.00046;
 Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1294 TCCGAGCTACTCGGAGCTGAGCGAGAGATGCGGTAC 1335
 ||||||||||||||||||||||||||||||||||||||||
 Db 84 TCCGAGCTACTCGGAGCTGAGCGAGAGATGCGGTAC 43
 RESULT 6
 A1932967/c 92 bp mRNA EST 02-SEP-1999
 LOCUS w040f09.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2457833 3'
 DEFINITION mRNA sequence.
 ACCESSION A1932967
 VERSION A1932967.1 GI:5671704
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 92)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL
 COMMENT On Dec 20, 1995 this sequence version replaced gi:1134104.

CONTACT: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/dbfp/image/image.html
 Seq primer: -40UP from Gibco.
 Location/Qualifiers
 1..92
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2457833"
 /clone_lib="NCI_CGAP_Gas4"
 /tissue_type="poorly differentiated adenocarcinoma with
 signed ring cell features"
 /lab_host="DH10B"
 /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"
 BASE COUNT 29 a 19 c 14 g 30 t
 ORIGIN
 Query Match 0.1%; Score 41; DB 62; Length 92;
 Best Local Similarity 100.0%; Pred. No. 0.001;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1392 GACAGAGCGAGACTCGCTCAAAAAAAAAAAAAAAAA 1432
 ||||||||||||||||||||||||||||||||||||||||
 Db 41 GACAGAGCGAGACTCGCTCAAAAAAAAAAAAAAAAA 1
 RESULT 7
 A0626158/c 104 bp DNA GSS 16-JUN-1999
 LOCUS CITBI-EI-2650P18.TR CITBI-EI Homo sapiens genomic clone 2650P18,
 DEFINITION genomic survey sequence.
 ACCESSION A0626158
 VERSION A0626158.1 GI:5088550
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 104)
 Zhao, S., Adams, M.D., Nierman, W., Malek, J., Shizuya, H., Simon, M. and
 Venter, J.C.
 Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
 Map Building
 Unpublished (1997)
 Other_GSSs: CITBI-EI-2650P18.TR
 Contact: Shaying Zhao, William Nierman, Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: hbe@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/tdb/hunggen/bac_end_search/bac_end_search.html.
 Seq primer: M13 Reverse
 Class: BAC ends
 Location/Qualifiers
 1..104
 /organism="Homo sapiens"

```
/db_xref="taxon:9606"
/clone="2650P18"
/clone_id="CTBI-E1"
/sex="male"
/cell_type="sperm"
/notes="Vector: pBelOAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library P"
BASE COUNT      12 a      28 c      22 g      42 t
ORIGIN

Query Match      0.1%; Score 41; DB 105; Length 104;
Best Local Similarity 100.0%; Pred. No. 0.00098;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1392 GACAGAGCAGAGCTCCGTCACAAAAA 1432
|||||
Db 94 GACAGAGCAGAGCTCCGTCACAAAAA 54

RESULT 8
D25879      52 bp      mRNA      EST      30-NOV-1995
LOCUS      HMGS05672 Human colon mucosa Homo sapiens cDNA clone cm2335 3',
DEFINITION      mRNA sequence.
ACCESSION      D25879
VERSION      D25879.1 GI:500543
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 52)
AUTHORS      Okubo,K., Yoshii,J., Yokouchi,H., Kameyama,M. and Matsubara,K.
TITLE      Global analysis of gene expression in colon mucosa: a large scale
random cDNA sequencing analysis
JOURNAL      Unpublished (1994)
COMMENT      Contact: Okubo,K., Itoh,K., Yoshii,J., Yokouchi,H. and Matsubara,K.
Institute for Molecular and Cellular Biology
Osaka University
3-1 Yamada-oka, Suita, Osaka 565, Japan.
FEATURES
source      1..52
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="cm2335"
/clone_id="Human colon mucosa"
/notes="Adult male, tissue type = colon mucosa"
BASE COUNT      12 a      17 c      14 g      9 t
ORIGIN

Query Match      0.1%; Score 40; DB 20; Length 52;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1376 GCACTCCAGCTGGGTGACAGAGCAGACTCGCTCAAA 1415
|||||
Db 13 GCACTCCAGCTGGGTGACAGAGCAGACTCGCTCAAA 52

RESULT 9
B48914      103 bp      DNA      GSS      08-APR-1999
LOCUS      RPCI11-4A12.TP RPCI-11 Homo sapiens genomic clone RPCI-11-4A12,
DEFINITION      genomic survey sequence.
ACCESSION      B48914
VERSION      B48914.1 GI:2601151
KEYWORDS      GSS.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
```

```
REFERENCE      1 (bases 1 to 103)
AUTHORS      Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
TITLE      Use of BAC End Sequences for Sequence-ready Map Building
JOURNAL      Unpublished (1997)
COMMENT      Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: madams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://BACPAC.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html
Seq primer: SP6
Class: BAC ends.
FEATURES
source      1..103
location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:750163"
/db_xref="taxon:9606"
/clone="RPCI-11-4A12"
/clone_id="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT      30 a      28 c      30 g      15 t
ORIGIN

Query Match      0.1%; Score 40; DB 81; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1168 TTGGGAGCCGAGCGCGGATCAGAGTCAAGAGAT 1207
|||||
Db 14 TTGGGAGCCGAGCGCGGATCAGAGTCAAGAGAT 53

RESULT 10
A1864984      90 bp      mRNA      EST      30-AUG-1999
LOCUS      WK06c11.x1 NCI-CGAP_Lym12 Homo sapiens cDNA clone IMAGE:2411540 3',
DEFINITION      similar to contains Alu repetitive element; mRNA sequence.
ACCESSION      A1864984
VERSION      A1864984.1 GI:5529091
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 90)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      On May 18, 1998 this sequence version replaced gi:3337794.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Life Technologies catalog #: 11547-015
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
www-bio.llnl.gov/bdrip/image/image.html
Seq primer: -40UP from G1bco.
```

Mon Apr 3 08:24:34 2000

us-08-852-495c-1-copy_90000_120000.1st

Page 6

FEATURES
source
1. 90
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2411540"
/clone_id="NCI-CGAP-Lym12"
/tissue_type="lymphoma, follicular mixed small and large cell"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: PCMV-SPORE6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: 0190 dt. Average insert size 1.25 kb. Life Technologies catalog #: 11547-015"
21 t
29 c
13 a
BASE COUNT
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 0.025; Length 90;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1337 CGGAGGCGGAGCTGCGAGCGGAGCGGCGCA 1373
DB 79 CGGAGGCGGAGCTGCGAGCGGAGCGGCGCA 43

RESULT 11
LOCUS
DEFINITION
AI801185 35 bp mRNA EST 06-JUL-1999
to79h04.x1 NCI-CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2184535 3'
similar to gb|U87921|HUMALC120 Human carcinoma cell-derived Alu
RNA transcript, (rRNA) contains element TAR1 repetitive element ;,
mRNA sequence.
AI801185
AI801185.1 GI:5366657
EST.
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 35)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Mar 16 1998 this sequence version replaced gi:2961671.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Koskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNM at:
www-bio.lnl.gov/bdrip/image/image.html

JOURNAL
COMMENT
CONTACT
TELEPHONE
EMAIL
TISSUE
CDNA
DNA
CLONE
WWW

FEATURES
source
1. 35
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="2184535"
/clone_id="NCI-CGAP_Gas4"
/tissue_type="poorly differentiated adenocarcinoma with
signed ring cell features"
/lab_host="DH10B"
/note="Organ: stomach; Vector: PCMV-SPORE6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: 0190 dt.
Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"

BASE COUNT
ORIGIN
5 a 10 c 10 g 10 t

Query Match
Best Local Similarity 100.0%; Pred. No. 0.2; Length 35;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1382 CAGCTGGGTGACAGAGCGGAGCTCGCTCAAAA 1416
DB 35 CAGCTGGGTGACAGAGCGGAGCTCGCTCAAAA 1

RESULT 12
LOCUS
DEFINITION
AA809831 88 bp mRNA EST 19-FEB-1998
oa40fil.s1 NCI-CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1307469 3'
similar to contains Alu repetitive element; , mRNA sequence.
AA809831
AA809831.1 GI:2879237
EST.
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 88)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Jan 19 1998 this sequence version replaced gi:2287025.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., David Allman,
Ph.D., Gerald Marti, M.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNM at:
www-bio.lnl.gov/bdrip/image/image.html

JOURNAL
COMMENT
CONTACT
TELEPHONE
EMAIL
TISSUE
CDNA
DNA
CLONE
WWW

FEATURES
source
1. 88
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1307469"
/clone_id="NCI-CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand CDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, 19D-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). CDNA synthesis was
primed with a Not I - 0190(dt) primer
15'-TGTACCAATCGAAGTGGAGCGGCGCGCTCATTTTCTTTTCTTTT-
3'. Double-stranded CDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT
ORIGIN
14 a 28 c 23 g 23 t

Query Match
0.1%; Score 35; DB 38; Length 88;

Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1173 GAGCCGAGCGCGGATCAGAGTACAGAGAT 1207

Db 62 GAGCGCGAGCGCGGATCAGAGTACAGAGAT 28

RESULT 13

R67088

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1995)
On Apr 5, 1995 this sequence version replaced gi:760768.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
Insert Size: 1093
High quality sequence stops: 65
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1093 Std Error: 0.00
Seq primer: Promega -21m13
High quality sequence stop: 65.

FEATURES

source

1.96

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:140793"

/clone_1lb="Soares placenta Nb2HP"

/sex="Female"

/dev_stage="Placenta obtained at birth (full term)"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: placenta; Vector: pTZ19 (Pharmacia) with a
modified polylinker; Site: 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
AAGTGAAGATTCGGCGCGAGGATTTTCTTTTCTTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTZ19 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Petina Bonaldo."

BASE COUNT 25 a 24 c 21 g 25 t 1 others
ORIGIN

Query Match 0.1%; Score 35; DB 22; Length 96;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3853 CCGCGCTCGGCTCCCAAGTGTGGATTACAG 3887

Db 1 CCGCGCTCGGCTCCCAAGTGTGGATTACAG 35

RESULT 14

A1591270

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)
On Mar 20, 1998 this sequence version replaced gi:2980211.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Herbert Morse, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bhrp/image/image.html
Seq primer: -40UP from Gibco.

FEATURES

source

1.102

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2246670"

/clone_1lb="NCI-CGAP_HSC3"

/tissue_type="CD34+, T negative, patient with chronic
myelogenous leukemia"

/lab_host="DH10B"

/note="Organ: bone marrow; Vector: pAMP1; mRNA made from
lymphoid tissue; cDNA made by oligo-dT priming.
Directionally cloned. Size selected on agarose gel,
average insert size 500 bp. Primary library,
non-amplified. cDNA library Preparation: David B.
Krizman, Ph.D. Reference: Krizman et al. (1996) Cancer
Research 56:5380-5383."

BASE COUNT 51 a 19 c 20 g 12 t
ORIGIN

Query Match 0.1%; Score 35; DB 48; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18986 GTGAACCCGCTCTACTTAAAAA 19020

Db 44 GTGAACCCGCTCTACTTAAAAA 78

RESULT 15

A0535244/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Unpublished (1997)
On Mar 20, 1998 this sequence version replaced gi:2980211.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Herbert Morse, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bhrp/image/image.html
Seq primer: -40UP from Gibco.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 103)
Zhao, S., Adams, M.D., Niernan, W., Malek, J., de Jong, P. and
Venter, J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
Unpublished (1997)
Contact: Shaying Zhao, William Niernan, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@jng.med.bufo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.bufo.edu/ordering>) or from
Research Genet cs (http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html).
Seq primer: T7
Class: BAC ends.

FEATURES
SOURCE

location/Qualifiers
1..103
/organism="Homo sapiens"
/db_xref="GDB:7621533"
/db_xref="taxon:9606"
/clone="RPCI-11-317H22"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT 31 a 27 c 27 g 18 t
ORIGIN

Query Match 0.1%; Score 35; DB 104; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5307 CCAGCTGCTCTGCACTCTGACCTCAAGTATC 5341
DB 63 CCAGCTGCTCTGCACTCTGACCTCAAGTATC 29

Search completed: March 30, 2000, 13:20:07
Job time: 122661 sec

Mon Apr 3 08:26:56 2000

us-09-425-665-1.rst

Page 11



1
2
3

4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 17:03:40 ; Search time 13413.1 Seconds

(without alignments)
-6791.452 Million cell updates/sec

Title: US-08-852-495C-1_COPY_115000_145000

Sequence: 1 AAATCTTCATTTTGTATA.....AGAAATTAACCTGATTAAA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 821193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : GenEmb1:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ov:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*
16: gb_ov:*
17: gb_ov:*
18: gb_ov:*
19: gb_ov:*
20: gb_ov:*
21: gb_ov:*
22: gb_ov:*
23: gb_ov:*
24: gb_ov:*
25: gb_ov:*
26: gb_ov:*
27: gb_ov:*
28: gb_ov:*
29: gb_ov:*
30: gb_ov:*
31: gb_ov:*
32: gb_ov:*
33: gb_ov:*
34: gb_ov:*
35: gb_ov:*
36: gb_ov:*
37: gb_ov:*
38: gb_ov:*
39: gb_ov:*
40: gb_ov:*
41: gb_ov:*
42: gb_ov:*
43: gb_ov:*

44: gb_htg6:*
45: em_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_htg5:*
50: gb_pl3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	0.1	60	5	AR051487
2	38	0.1	76	5	AR051499
3	36	0.1	67	9	HUMALUNCD
4	35	0.1	51	9	HUMD6B03M3
5	35	0.1	79	13	HS024138
6	34	0.1	66	40	AF087511
7	33	0.1	79	10	S73203
8	32	0.1	80	9	HUMBRKFAA
9	32	0.1	80	9	HUMBRKFAA
10	31	0.1	60	11	HSCBEB5115
11	31	0.1	76	11	AF032238
12	31	0.1	97	9	HUMDLRA2
13	31	0.1	97	9	HUMDLRDLJ
14	31	0.1	103	9	HUMALCE221
15	30	0.1	80	9	HUMBRKFAA
16	30	0.1	80	9	HUMBRKFAA
17	30	0.1	95	9	HSTHERKIB
18	30	0.1	100	9	HS1AS27
19	30	0.1	101	10	S79560
20	30	0.1	101	10	S79561
21	29	0.1	90	9	HUMDLRFL
22	29	0.1	102	13	G32906
23	29	0.1	102	13	G37854
24	29	0.1	104	13	G20851
25	28	0.1	50	5	I23510
26	28	0.1	50	5	I28359
27	28	0.1	50	5	I28514
28	28	0.1	50	5	I41125
29	28	0.1	50	5	I49056
30	28	0.1	50	5	I70295
31	28	0.1	50	5	I90068
32	28	0.1	57	5	A45333
33	28	0.1	57	5	AR061138
34	28	0.1	57	9	HSDNAPYTB
35	28	0.1	50	9	HUMDLRPM
36	28	0.1	103	13	HSBIC8R
37	27	0.1	40	5	A68621
38	27	0.1	63	13	HUMUT5302A
39	27	0.1	67	9	HUMALUNCD
40	27	0.1	71	13	HUMUT41A
41	27	0.1	72	13	HUMUT1667A
42	27	0.1	73	5	A08915
43	27	0.1	80	34	HUMBRKFAA
44	27	0.1	80	34	DRORS1584
45	27	0.1	80	34	DRORS1586

ALIGNMENTS

RESULT 1
LOCUS AR051487 60 bp DNA
DEFINITION Sequence 57 from patent US 5830670.
ACCESSION AR051487
VERSION AR051487.1 GI:5974851

29-SEP-1999

KEYWORDS .
SOURCE Unknown
ORGANISM Unknown
REFERENCE 1 (bases 1 to 60)
AUTHORS de la Monte, S. and Wands, J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 57 03-NOV-1998;
FEATURES Location/Qualifiers
SOURCE 1. 60
BASE COUNT 12 a 14 c 15 g 19 t
ORIGIN

Query Match 0.1%; Score 38; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17232 TTTTGTATTTTAGTAGAGACAGGTTTCACCGTGT 17269
|||||
Db 16 TTTTGTATTTTAGTAGAGACAGGTTTCACCGTGT 53

RESULT 2
LOCUS AR051499 76 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 69 from patent US 5830670.
ACCESSION AR051499
VERSION AR051499.1 GI:5974863
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 76)
AUTHORS de la Monte, S. and Wands, J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 69 03-NOV-1998;
FEATURES Location/Qualifiers
SOURCE 1. 76
BASE COUNT 15 a 19 c 18 g 24 t
ORIGIN

Query Match 0.1%; Score 38; DB 5; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17232 TTTTGTATTTTAGTAGAGACAGGTTTCACCGTGT 17269
|||||
Db 16 TTTTGTATTTTAGTAGAGACAGGTTTCACCGTGT 53

RESULT 3
LOCUS HUMALUAND 67 bp DNA PRI 08-OCT-1994
DEFINITION Homo sapiens 4000 year old remains from Nekht-anKh Alu repeat fragment 12:5.
ACCESSION L36843
VERSION L36843.1 GI:556196
KEYWORDS Alu repeat.
SOURCE Homo sapiens (individual isolate 4000 year old remains from Nekht-anKh) liver DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eukaryota; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Paabo, S.
TITLE Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (6), 1939-1943 (1989)
MEDLINE 89184542
FEATURES Location/Qualifiers
SOURCE 1. 67
/organism="Homo sapiens"
/isolate="4000 year old remains from Nekht-anKh"
/db_xref="taxon:9606"
/tissue_type="liver"
repeat_region
/rpl_family="Alu"
BASE COUNT 15 a 18 c 23 g 11 t
ORIGIN

Query Match 0.1%; Score 36; DB 9; Length 67;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22232 GTGGCTCACACCTGTAATCCAGACATTGGAGGC 22267
|||||
Db 9 GTGGCTCACACCTGTAATCCAGACATTGGAGGC 44

RESULT 4
LOCUS HUMD6B03M3/C 51 bp mRNA PRI 04-FEB-1999
DEFINITION Human HepG2 3' region MDOI cDNA, clone hmd6b03m3.
ACCESSION D17279
VERSION D17279.1 GI:598922
KEYWORDS gene signature
SOURCE Homo sapiens Male cell_line:HepG2 cDNA to mRNA, clone_1lb:kiserv.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 51)
AUTHORS Matoba, R.
TITLE Direct Submission
JOURNAL Submitted (21-JUL-1993) to the DDBJ/EMBL/GenBank databases. Ryo Matoba, Osaka University, Institute for Molecular and Cellular Bio; 1-3, Yamada-oka, Suita, Osaka 565, Japan (E-mail:matoba@inherit.imcb.osaka-u.ac.jp; Tel:81-6-877-5111(ex.3314), Fax:81-6-877-1922)
REFERENCE 2 (bases 1 to 51)
AUTHORS Matoba, R., Okubo, K., Hori, N., Fukushima, A. and Matsubara, K.
TITLE Improves analysis of 5'-coding information to a 3'-directed cDNA library
JOURNAL Gene 146 (2), 199-207 (1994)
MEDLINE 94357437
COMMENT Submitted (21-Jul-1993) to DDBJ by: Ryo Matoba
Research Institute of Innovative Technology for the Earth 9-2 Kirugadai Kita-kyo, Soraku-gun, Kyoto Japan, 619-02 Soraku-gun, Kyoto Japan, 619-02
Phone: 07747-5-2308
Fax: 07747-5-2321.
FEATURES Location/Qualifiers
SOURCE 1. 51
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HepG2"
/clone_1lb="Kiserv"
/sex="Male"
BASE COUNT 13 a 19 c 12 g 7 t
ORIGIN

Query Match 0.1%; Score 35; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 6.8e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16302 CTTGTGCGCCAGGCTGAGTGCAGTGCGGTATC 16336

```

Db      35  CTTCTGCCCGAGCTGAGATGCAGTGGCTGATC 1
|||||
RESULT  5
LOCUS    HSU24138      79 bp      DNA
DEFINITION Hylobates sp. sequence homologous to Human dinucleotide repeat
ACCESSION U24138
VERSION   U24138.1 GI:773649
KEYWORDS  STS.
SOURCE    gibbon.
ORGANISM  Hylobates sp.
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS   Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
TITLE     1 (bases 1 to 79)
JOURNAL   Arcot,S.S., Wang,Z., Weber,J.L., Deininger,P.L. and Batzer,M.A.
MEDLINE   Alu repeats: a source for the genesis of primate microsatellites
REFERENCE 2 Genomics 29 (1), 136-144 (1995)
AUTHORS   96079101
TITLE     2 (bases 1 to 79)
JOURNAL   Batzer,M.A., Rubin,C.M., Hellmann-Blumberg,U., Alegria-Hartman,M.,
MEDLINE   Leeftang,E.P., Stern,J.D., Bazan,H.A., Shaikh,T.H., Deininger,P.L.
REFERENCE 3 Dispersion and insertion polymorphism in two small subfamilies of
AUTHORS   recently amplified human Alu repeats
TITLE     J. Mol. Biol. 247 (3), 418-427 (1995)
JOURNAL   95230683
MEDLINE   3 (bases 1 to 79)
AUTHORS   Weber,J.L.
TITLE     Direct Submision
JOURNAL   Submitted (30-MAR-1995) James L. Weber, Medical Genetics,
MEDLINE   Marshfield Medical Research Foundation, 1000 N. Oak Ave.,
REFERENCE 1 Marshfield, WI 54449, USA
AUTHORS   Location/Qualifiers
TITLE     1. 79
JOURNAL   /organism="Hylobates sp."
MEDLINE   /db_xref="taxon:9581"
AUTHORS   /chromosome="7"
JOURNAL   /note="Human Mtd158 homolog, GenBank Accession Number
MEDLINE   L15398"
AUTHORS   30 a 17 c 15 g 17 t
BASE COUNT
ORIGIN
Query Match 0.1%; Score 35; DB 13; Length 79;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 22361 GCGGTGTCGATGCTGCTATCCAGCTACTC 22395
|||||
Db      45  GCGGTGTCGATGCTGCTATCCAGCTACTC 79
|||||
RESULT  6
LOCUS    AF087511      66 bp      mRNA
DEFINITION Homo sapiens clone ENAC+22 epithelial sodium channel alpha subunit
ACCESSION AF087511
VERSION   AF087511.1 GI:5870626
KEYWORDS
SOURCE    human.
ORGANISM  Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS   Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE     1 (bases 1 to 66)
JOURNAL   Oh,Y. and Warnock,D.G.
MEDLINE   An Alu cassette in the human epithelial sodium channel
REFERENCE 2 unpublished
AUTHORS   2 (bases 1 to 66)
JOURNAL   Oh,Y. and Warnock,D.G.

```

```

TITLE     Direct Submission
JOURNAL   Submitted (26-AUG-1998) Medicine, UAB, 1720 7th Ave. So.,
AUTHORS   Birmingham, AL 35294, USA
FEATURES  Location/Qualifiers
SOURCE    1. 66
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
          /chromosome="12"
          /clone="ENAC+22"
          <1..>66
          /gene="SCNN1A"
          /note="alternatively spliced epithelial sodium channel
          alpha subunit containing intronic Alu sequence"
          <1..>66
          /gene="SCNN1A"
BASE COUNT 12 a 20 c 24 g 10 t
ORIGIN
Query Match 0.1%; Score 34; DB 40; Length 66;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 22235 GCTCACACCTGTATCCACAGCTTGGAGGCC 22268
|||||
Db      14  GCTCACACCTGTATCCACAGCTTGGAGGCC 47
|||||
RESULT  7
LOCUS    S73203      79 bp      DNA
DEFINITION ALL-1 (tandem duplication) [human, acute myeloid leukemia patient,
ACCESSION S73203
VERSION   S73203.1 GI:685048
KEYWORDS
SOURCE    human acute myeloid leukemia patient.
ORGANISM  Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS   Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE     1 (bases 1 to 79)
JOURNAL   Schichman,S.A., Calligaris,M.A., Strout,M.P., Carter,S.L., Gu,Y.,
MEDLINE   Canaan,I.E., Bloomfield,C.D. and Croce,C.M.
REFERENCE 2 ALL-1 tandem duplication in acute myeloid leukemia with a normal
AUTHORS   karyotype involves homologous recombination between Alu elements
JOURNAL   Cancer Res. 54 (16), 4277-4280 (1994)
MEDLINE   94320053
REFERENCE 3 Genbank staff at the National Library of Medicine created this
AUTHORS   entry [NCBI gblseq 155288] from the original journal article.
JOURNAL   This sequence comes from Fig. 3a.
MEDLINE   Map location: 11.
REFERENCE 4 Location/Qualifiers
AUTHORS   1. 79
JOURNAL   /organism="Homo sapiens"
MEDLINE   /db_xref="taxon:9606"
AUTHORS   1. 79
JOURNAL   /partial
AUTHORS   /gene="ALL-1"
BASE COUNT 19 a 18 c 29 g 13 t
ORIGIN
Query Match 0.1%; Score 33; DB 10; Length 79;
Best Local Similarity 100.0%; Pred. No. 6.6e-05;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 22418 GCTTGAACCCAGAGGCGGAGGTTCAGTGGAGC 22450
|||||
Db      37  GCTTGAACCCAGAGGCGGAGGTTCAGTGGAGC 69
|||||
RESULT  8
JOURNAL   HUMBRKFAA/c

```

LOCUS HUMBRKFA 80 bp DNA PRI 27-APR-1993
DEFINITION Human alpha-galactosidase breakpoint region.
ACCESSION M36131.1 GI:179541
VERSION M36131.1
KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
AUTHORS Kornreich,R., Bishop,D.F. and Desnick,R.J.
TITLE Alpha-galactosidase A gene rearrangements causing Fabry disease:
Identification of short direct repeats at breakpoints in an
Alu-rich gene
JOURNAL J. Biol. Chem. 265, 9319-9326 (1990)
MEDLINE 90264427
FEATURES
Location/Qualifiers
1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 19 a 28 c 20 g 13 t
ORIGIN

Query Match 0.1%; Score 32; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 32; Conservative 0; Indels 0; Gaps 0;

QY 22242 CCTGTAATCCGACACTTTGGAGGCCAAGT 22273
DB 47 CCTGTAATCCGACACTTTGGAGGCCAAGT 16

RESULT 9
LOCUS HUMBRKFA 80 bp DNA PRI 27-APR-1993
DEFINITION Human alpha-galactosidase breakpoint family F.
ACCESSION M36132
VERSION M36132.1 GI:179542
KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
AUTHORS Kornreich,R., Bishop,D.F. and Desnick,R.J.
TITLE Alpha-galactosidase A gene rearrangements causing Fabry disease:
Identification of short direct repeats at breakpoints in an
Alu-rich gene
JOURNAL J. Biol. Chem. 265, 9319-9326 (1990)
MEDLINE 90264427
FEATURES
Location/Qualifiers
1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 17 a 29 c 16 g 18 t
ORIGIN

Query Match 0.1%; Score 32; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 32; Conservative 0; Indels 0; Gaps 0;

QY 22242 CCTGTAATCCGACACTTTGGAGGCCAAGT 22273
DB 47 CCTGTAATCCGACACTTTGGAGGCCAAGT 16

RESULT 10
LOCUS HSCBRB515/c 60 bp DNA PRI 26-SEP-1998
DEFINITION Homo sapiens core-binding factor beta subunit (CBFB) gene, intron
5, partial sequence.

ACCESSION AF084969
VERSION AF084969.1 GI:3659831
KEYWORDS 15 of 16
SEGMENT
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
AUTHORS van der Reijden,B.A., Dauwerse,H.G., Gilles,R.H.,
Jagmoohan-Changur,S., Wijnen,G.C., Liu,P.P., Smit,B., Wessels,H.W.,
Beverloek,G.C., Jotterand-Bellomo,M., Martinet,D.,
Chlematier,D.M.F., Lafage-Pochitaloff,M., Gabert,J., Reiffers,J.,
Bilhou-Nabera,C.E., van Ommen,G.-J.B., Hagemeijer,A. and
Breuning,M.H.
TITLE Genomic acute myeloid leukemia-associated inv(16)(p13q22)
breakpoints are tightly clustered
JOURNAL Oncogene (1998) in press
AUTHORS van der Reijden,B.A., Dauwerse,H.G., Gilles,R.H.,
Jagmoohan-Changur,S. and Breuning,M.H.
TITLE Nonoverlapping genomic inv(16)(p13q22) CBFB intron 5 breakpoint
regions ordered centromeric to telomeric
JOURNAL unpublished
AUTHORS van der Reijden,B.A., Dauwerse,H.G., Gilles,R.H.,
Jagmoohan-Changur,S. and Breuning,M.H.
TITLE Direct Submission
JOURNAL Submitted (20-AUG-1998) Hematology, Erasmus University Rotterdam,
Dr. Molwaterplein 50, Rotterdam 3015 GE, The Netherlands
FEATURES
Location/Qualifiers
1..60
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q22"
BASE COUNT 13 a 20 c 15 g 12 t
ORIGIN

Query Match 0.1%; Score 31; DB 11; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23887 TCATGCTGTATCCGACACTTTGGAGGC 23917
DB 39 TCATGCTGTATCCGACACTTTGGAGGC 9

RESULT 11
LOCUS AF032238/c 76 bp DNA PRI 13-OCT-1998
DEFINITION Otlemur crassicaudatus clone GALI17 Galago Alu.
ACCESSION AF032238
VERSION AF032238.1 GI:2642069
KEYWORDS
SOURCE
ORGANISM Otlemur crassicaudatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Strepsirhini; Galagonidae; Otlemur.
REFERENCE
AUTHORS Zietkiewicz,E., Richer,C., Smet,D. and Labuda,D.
TITLE Monophyletic origin of Alu elements in primates
JOURNAL J. Mol. Evol. 47 (2), 172-182 (1998)
MEDLINE 98360010
AUTHORS Zietkiewicz,E., Richer,C., Smet,D. and Labuda,D.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-1997) Pediatrics, University of Montreal,
Sainte-Justine Hospital, Research Center (Charles Bruneau Center of
Cancerology), 3175 Cote Sainte-Catherine, Montreal, Quebec H3T-1C5,
Canada
COMMENT Sequences corresponding to the PCR primers used to amplify

Alu element are not included.

Forward primer: 5'-ggcgcggtggtcaccg-3';

Reverse primer: 5'-ttttttgagcagcagtcctc-3'.

FEATURES
source
Location/Qualifiers
1..76
/organism="Otolemur crassicaudatus"
/db_xref="taxon:9463"
/clone="GAL1127"
/note="PCR-amplified"

repeat_region
1..76
/rpt_family="Galago Alu"
/rpt_type="dispersed"

BASE COUNT
20 a 20 c 21 g 15 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Score 31; DB 11; Length 76;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7702 TGGTCTCAACTCCTGAGCTCAAGCAATCCA 7732
|||||
Db 66 TGGTCTCAACTCCTGAGCTCAAGCAATCCA 36

RESULT 12
HUMDLRA2/c
LOCUS
DEFINITION Human low density lipoprotein receptor gene (LDLR), Intron 5
(partial).
ACCESSION M14180.1 GI:187098
VERSION low density lipoprotein receptor-1.
KEYWORDS 2 of 2
SEGMENT
SOURCE Human white blood cell DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Hobbs, H.H., Brown, M.S., Goldstein, J.L. and Russell, D.W.
TITLE Deletion of exon encoding cysteine-rich repeat of low density lipoprotein receptor alters its binding specificity in a subject with familial hypercholesterolemia
J. Biol. Chem. 261 (28), 13114-13120 (1986)
JOURNAL MEDLINE 87008518
COMMENT Analysis of the LDL-receptor gene of a patient with familial hypercholesterolemia (FH) revealed the deletion of exon 5 resulting from a homologous recombination between repetitive Alu sequences of Intron 4 and Intron 5.

FEATURES
source
Location/Qualifiers
1..97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19p13.3"
join(M14178.1:1..97,1..97)
/gene="LDLR"
<1..>97
/gene="LDLR"
/note="LDLR Intron E; G00-119-362"
misc_feature
42..72
/gene="LDLR"
/note="deletion target sequence"

BASE COUNT
19 a 34 c 25 g 19 t
ORIGIN Chromosome 19p13.2-p13.1.

Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 97;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22242 CCTGTATCCACACTTTGGAGCCCAAG 22272
|||||
Db 77 CCTGTATCCACACTTTGGAGCCCAAG 47

RESULT 13
HUMDLRDJ/c
LOCUS
DEFINITION Human familial hypercholesterolemia 626-a gene with a deletion of exon 5.
ACCESSION M14179.1 GI:187101
VERSION low density lipoprotein receptor-1; very low density lipoprotein. Human fibroblast DNA, from patient FH-626.
KEYWORDS
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Hobbs, H.H., Brown, M.S., Goldstein, J.L. and Russell, D.W.
TITLE Deletion of exon encoding cysteine-rich repeat of low density lipoprotein receptor alters its binding specificity in a subject with familial hypercholesterolemia
J. Biol. Chem. 261, 13114-13120 (1986)
JOURNAL MEDLINE 87008518
COMMENT Analysis of the LDL-receptor gene of a patient with familial hypercholesterolemia (FH) revealed the deletion of exon 5 resulting from a homologous recombination between repetitive Alu sequences of Intron 4 and Intron 5.

FEATURES
source
Location/Qualifiers
1..97
/organism="Homo sapiens"
/db_xref="taxon:9606"

misc_feature
42..72
/note="deletion target sequence"
BASE COUNT
22 a 34 c 23 g 18 t
ORIGIN Chromosome 19p13.2-p13.1.

Query Match
Best Local Similarity 100.0%; Score 31; DB 9; Length 97;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22242 CCTGTATCCACACTTTGGAGCCCAAG 22272
|||||
Db 77 CCTGTATCCACACTTTGGAGCCCAAG 47

RESULT 14
HUMALCE221/c
LOCUS
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Sinnett, D., Richer, C., Deragon, J.-M. and Labuda, D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
J. Mol. Biol. (1992) In press
JOURNAL
FEATURES
source
Location/Qualifiers
1..103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="NTera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"

BASE COUNT
25 a 27 c 33 g 18 t
ORIGIN

Mon Apr 3 08:24:03 2000

us-08-852-495c-1_copy_115000_145000.rge

Page 6

Query Match 0.1%; Score 31; DB 9; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17128 ATCTGCGCTCACTGCAACCTCCGCTCCCG 17158
|||||
DB 85 ATCTGCGCTCACTGCAACCTCCGCTCCCG 55

RESULT 15

HUMBRKFAA 80 bp DNA PRI 27-APR-1993
LOCUS Human alpha-galactosidase breakpoint region.
DEFINITION M36131
ACCESSION M36131.1 GI:179541
VERSION M36131.1
KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 80)
AUTHORS Kornreich, R., Bishop, D.F. and Desnick, R.J.
TITLE Alpha-galactosidase A gene rearrangements causing Fabry disease:
Identification of short direct repeats at breakpoints in an
Alu-rich gene

JOURNAL J. Biol. Chem. 265, 9319-9326 (1990)
MEDLINE 90264427
FEATURES Location/Qualifiers

source 1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 19 a 28 c 20 g 13 t
ORIGIN

Query Match 0.1%; Score 30; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16511 GCCCACCTTGGCTCCCAAGTGTGGGAT 16540
|||||
DB 12 GCCCACCTTGGCTCCCAAGTGTGGGAT 41

Search completed: March 30, 2000, 23:41:00
Job time: 159409 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 17:38:54 ; Search time 612.33 Seconds
(Without alignments)
12258.117 Million cell updates/sec

Title: US-08-852-495c-1_COPY_115000_145000
Perfect score: 30001
Sequence: 1 AATCTTCACATTTTGTAATA.....AGAAATAACCTGATTTAA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 311585 seqs, 125096042 residues

Word size : 0

Total number of hits satisfying chosen parameters: 433264

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	28	0.1	40	1 Q55168	Sequence of primer
C 2	28	0.1	48	1 T04081	Trypsin inhibitory
C 3	28	0.1	51	1 T25038	Human gene signatu
C 4	28	0.1	57	1 Q95177	Simple tandem repe
C 5	28	0.1	60	1 T69401	PCR primer A11. Nu
C 6	28	0.1	86	1 Q11760	Self-complementary
C 7	28	0.1	98	1 X00174	Porcine reproducti
C 8	27	0.1	40	1 V19044	Alu PCR primer 1.
C 9	27	0.1	61	1 T23178	Human gene signatu
C 10	27	0.1	80	1 V37197	Oligonucleotide se
C 11	27	0.1	88	1 V39744	Microsatellite ana
C 12	27	0.1	92	1 T88439	Padlock probe alph
C 13	27	0.1	95	1 Q75099	DNA molecule encod
C 14	27	0.1	100	1 V68827	Plasmid pOKSC18a c
C 15	26	0.1	27	1 V71935	Anchored poly T RT
C 16	26	0.1	32	1 Q27389	Inter-Alu specific
C 17	26	0.1	35	1 Q27391	Inter-Alu specific
C 18	26	0.1	40	1 V19045	Alu PCR primer 2.
C 19	26	0.1	48	1 Q86183	Primer SINKball1700
C 20	26	0.1	48	1 T30807	Sindbis PCR primer
C 21	26	0.1	48	1 T35073	Sindbis-based, tum
C 22	26	0.1	48	1 V42384	Reverse PCR primer
C 23	26	0.1	48	1 V70704	Reverse PCR primer
C 24	26	0.1	69	1 V02148	Human secreted pro
C 25	26	0.1	69	1 T88081	3' portion of cDNA
C 26	26	0.1	79	1 V26706	Human novel secret
C 27	26	0.1	79	1 V32414	Homo sapiens clone
C 28	26	0.1	85	1 V05720	Nucleotide sequenc
C 29	26	0.1	92	1 T39467	Growth regulatory
C 30	26	0.1	92	1 V37485	Human growth regul
C 31	26	0.1	99	1 T91300	Human M97-2 secret
C 32	26	0.1	101	1 V00420	3' fragment of c10
C 33	25	0.1	27	1 T93831	Phosphodiester o11

C 34	25	0.1	29	1 V15487
C 35	25	0.1	37	1 T28941
C 36	25	0.1	37	1 V12343
C 37	25	0.1	41	1 V03013
C 38	25	0.1	46	1 T65781
C 39	25	0.1	47	1 T98523
C 40	25	0.1	48	1 Q93960
C 41	25	0.1	50	1 T39266
C 42	25	0.1	50	1 T92388
C 43	25	0.1	50	1 V26819
C 44	25	0.1	50	1 V37933
C 45	25	0.1	50	1 X02643

ALIGNMENTS

RESULT 1	055168/c	055168 standard; DNA; 40 BP.	
ID	055168:		
AC	055168:		
DT	21-JUL-1994 (first entry)		
DE	Sequence of primer for PCR amplification of HIV-LP Pt.1 isolate		
KW	Human immunodeficiency virus; HIV-LP; PCR primer; ss.		
OS	Synthetic.		
PN	W09400562-A.		
PD	06-JAN-1994.		
PF	23-JUN-1993; U06162.		
PR	24-JUN-1992; US-903421.		
PA	(CORR) CORNELL RES FOUND INC.		
PA	(UANY) UNIV NEW YORK MT SINAI.		
PA	SCHOOL MEDICINE.		
PI	Gelman IH, Laurence JC;		
DR	WPI; 94-02620/03.		
PT	HIV-LP useful in vaccine formulations - 1s novel HIV virus		
PS	distinct from HIV-1 and or HIV-2 viruses		
CC	Example; page 7; 75pp; English.		
CC	HIV-LP is a new variant of the HIV family. A cDNA first strand was		
CC	synthesised from Pt. 1 pellet using M1V RT. The product was		
CC	converted into dsDNA and this cDNA was amplified by PCR using		
CC	primers 055167 and 055168.		
SQ	Sequence 40 BP; 3 A; 2 C; 3 G; 32 T;		

Query Match 0.1%; Score 28; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29696 AAAAAAAAAAAAAAAAAAAGAAAT 29723

Db 34 AAAAAAAAAAAAAAAAAAAGAAAT 7

RESULT 2
T04081/c
ID T04081 standard; DNA; 48 BP.
AC T04081:
DT 16-MAY-1996 (first entry)
DE Trypsin inhibitory protein cDNA antisense primer-1.
KW Trypsin; inhibitor; human T98G cells; pancreatitis; shock; DIC;
KM multiple organ failure; disseminated intravascular coagulation; ss.
OS Synthetic.
PN J07242700-A.
PD 19-SEP-1995.
PF 04-MAR-1994; 059906.
PR 04-MAR-1994; JP-059906.
PA (MOCH) MOCHIDA PHARM CO LTD.
DR WPI; 95-355285/46.
PT Polypeptide having trypsin inhibitory activity - for the treatment
of pancreatitis, shock, multi-organ failure, etc.
PS Example 7; Page 15; 25pp; Japanese.
CC A polypeptide having trypsin inhibitory activity, mol.wt. 22-28 kD
or 17-23 kD (as determined by SDS-PAGE under reducing or non-

CC reducing conditions, respectively) and an N-terminal sequence as in
CC R79913 is claimed. The polypeptide was isolated from human T98G
CC cells and is useful for treating pancreaticitis, shock, multiple
CC organ failure and disseminated intravascular coagulation. The
CC present sequence is that of an antisense primer used for cloning
CC T98G cDNA coding for the full-length polypeptide
SQ Sequence 48 Bp; 2 A; 1 C; 1 G; 44 T;

Query Match 0.1%; Score 28; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 29696 AAAAAAAAAAAAAAAAAAGAGT 29723
DB 32 AAAAAAAAAAAAAAAAAAGAGT 5

RESULT 3
ID T25038/c standard; cDNA to mRNA; 51 Bp.
AC T25038;
DT 11-NOV-1996 (first entry)
DE Human gene signature HOMG507164.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-41.
PD 01-JUN-1995
PF 11-NOV-1994; J01915
PR 12-NOV-1993; JP-355504.
PA (HABS/) MATSUBARA K.
PI (OKUBO/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1: Page 1755; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in 19901-126837 and which is able to hybridize to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridize with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 51 Bp; 14 A; 16 C; 12 G; 9 T;

Query Match 0.1%; Score 28; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16309 GCCCAGGCTGAGTGCAGTGCATGATC 16336
DB 28 GCCCAGGCTGAGTGCAGTGCATGATC 1

RESULT 4
ID 095177 standard; DNA; 57 Bp.
AC 095177;
DT 08-FEB-1996 (first entry)

DE Simple tandem repeat (STR) corresponding to the 2nd part of wglaz.
KW Simple tandem repeat; STR; wglaz; treatment; genetic; diagnosis;
KW characterisation; mapping; linkage studies; analysis; alleles;
KW second part; ss.
OS Synthetic.
PN WO9517522-42.
PD 29-JUN-1995
PF 21-DEC-1993; GB-026052.
PR (UYLE-) UNIV. LEICESTER.
PI Armour J, Jeffreys AJ;
DR WPI: 95-240682/31.
PT Identifying simple tandem repeat loci in DNA - by screening DNA
PT library to enrich for fragments contg. the repeats before cloning
PT and rescreening, also simple tandem repeats for treatment or
PT diagnosis

PS Claim 26: Page 17; 51pp; English.
CC 095177 is a simple tandem repeat (STR) corresponding to the 2nd part
CC of wglaz. The STR can be used for treatment and diagnosis in human
CC and veterinary medicine, partic. for genetic characterisation,
CC mapping, linkage studies and analysis/diagnosis of acquired disease
CC alleles.
SQ Sequence 57 Bp; 7 A; 11 C; 18 G; 21 T;

Query Match 0.1%; Score 28; DB 1; Length 57;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16309 GCCCAGGCTGAGTGCAGTGCATGATC 16336
DB 30 GCCCAGGCTGAGTGCAGTGCATGATC 57

RESULT 5
ID T69401/c standard; DNA; 60 Bp.
AC T69401;
DT 26-AUG-1997 (first entry)
DE PCR primer ALI.
KW Pheromone receptor; vomeronasal sensory neuron; social behaviour;
KW maternal behaviour; reproductive behaviour; fertility;
KW hormone secretion; polymerase chain reaction; PCR; primer; ss.
OS Synthetic.
PN WO9714790-A1.
PD 24-APR-1997.
PF 18-OCT-1996; U16637.
PR 19-OCT-1995; US-005698.
PA (UYCO) UNIV COLUMBIA NEW YORK.
PI Axel R, Dulac C;
DR WPI: 97-245107/22.
PT Nucleic acid molecule encoding vertebrate pheromone receptor -
PT useful to identify modulators for control of reproductive and social
PT behaviour, fertility and hormone secretion
PS Disclosure: Page 30; 133pp; English.
CC PCR primer ALI (T69401) was used to amplify cDNA derived from rat
CC main olfactory epithelium (MOE) and vomeronasal organ (VNO)
CC neurons. Data suggested that the amplified cDNA from individual
CC neurons contained an accurate representation of sequences in the
CC mRNA. Amplified cDNAs from single cells were used as probes,
CC leading to the isolation of a clone (T69545) coding for rat
CC pheromone receptor VNI (W19103).
SQ Sequence 60 Bp; 11 A; 8 C; 8 G; 33 T;

Query Match 0.1%; Score 28; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29696 AAAAAAAAAAAAAAAAAAGAGT 29723
DB 60 AAAAAAAAAAAAAAAAAAGAGT 33

```
RESULT 6
ID 011760/c
ID 011760 standard; DNA: 86 BP.
AC 011760:
DE 22-JUL-1991 (first entry)
DE Self-complementary, T7 promoter hairpin-forming sequence #2.
KW T7 bacteriophage; DNA-dependent RNA polymerase;
KW target sequence amplification; ss.
OS Synthetic.
FH Key
FH stem_loop
FH Location/Qualifiers
FT 1..86
FT /*tag= a
FT /note= "number of T's in loop is 50"
PN EP-427073-A.
PD 15-MAY-1991.
PF 27-OCT-1990; 120650.
PR 09-NOV-1989; US-434372.
PR 23-AUG-1990; US-569991.
PA (MOLE-) MOLECULAR DIAGNOSTI.
PI Dataagupta N.
PI WPI: 91-141638/20.
PT Nucleic acid probe for amplification and detection of target
PT sequence - capable of forming ligatable hairpin structured
PT promoter and transcribing target sequence, is sensitive and
PT useful in medical diagnosis
PS Claim 2, Page 11: 15pp; English.
CC The sequence is an example of a preferred T7 RNA polymerase
CC promoter for use in the invention. There can be from 2 to 50 T's in
CC the loop region. A probe sequence is ligated to the 3' end of the
CC promoter region. Upon hybridisation of the probe to a target sequence
CC and ligation of the hybridised target sequence to the 5' end of the
CC hairpin-forming sequence, the target sequence can be transcribed by
CC T7 RNA polymerase. This allows the target sequence to be amplified
CC (and detected) using a single oligonucleotide component (c.f. PCR).
CC See also 011759 and 011761-011764.
SQ Sequence 86 BP; 12 A; 6 C; 6 G; 62 T;
```

```
Query Match 0.1%; Score 28; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 29692 TATTAATAAAAAAAAAAAAAAAAAA 29719
DB 73 TATTAATAAAAAAAAAAAAAAAAAA 46
```

```
RESULT 7
ID X00174/c
ID X00174 standard; DNA: 98 BP.
AC X00174:
DE 23-MAR-1999 (first entry)
DE Porcine reproductive and respiratory syndrome virus PCR primer PRRSV-2.
KW Equine arteritis virus; EAV; vaccine; structural gene; PRRSV;
KW porcine reproductive and respiratory syndrome virus; recombinant virus;
KW PCR primer; ss.
OS Synthetic.
OS Porcine reproductive and respiratory syndrome virus.
PN WO9855626-A2.
PD 10-DEC-1998.
PF 05-JUN-1998; U12141.
PR 05-JUN-1997; US-048662.
PA (ORIG-) ORIGEN INC.
PI Consensus PM, Reilly JD, Spatz SJ;
PI WPI: 99-080829/07.
PT New recombinant porcine reproductive and respiratory syndrome virus
PT - containing nucleic acid encoding a polymerase from an RNA virus
PT and open reading frames 2-7 of the porcine virus, used particularly
PT in vaccines
PS Example 2; Page 21: 55pp; English.
CC The present invention describes a nucleic acid which encodes a
CC polymerase from an RNA virus, excluding porcine reproductive and
```

```
CC respiratory syndrome virus (PRRSV), and open reading frames (ORFs) 2-7
CC of PRRSV. The use of a polymerase gene from RNA viruses can provide for
CC production of less mutagenic recombinant viruses. The recombinant
CC viruses can be used in vaccines which have a reduced risk of loss or
CC reduction of efficacy. The vaccines are used particularly for
CC protecting swine against PRRSV. The high fidelity RNA polymerase gene
CC can be used as a marker that allows organisms vaccinated with such a
CC vaccine to be distinguished from organisms naturally infected with wild
CC type strains of virus or other vaccines. The present sequence represents
CC a PCR primer used in an example from the present invention.
SQ Sequence 98 BP; 11 A; 14 C; 8 G; 65 T;
```

```
Query Match 0.1%; Score 28; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 29696 AAAAAAAAAAAAAAAAAAGNAT 29723
DB 32 AAAAAAAAAAAAAAAAAAGNAT 5
```

```
RESULT 8
ID V19044
ID V19044 standard; DNA: 40 BP.
AC V19044:
DE 28-JUL-1998 (first entry)
DE Alu PCR primer 1.
KW PCR; primer; amplification; Alu repeat sequence; vector;
KW circular yeast artificial chromosome; YAC; ss.
OS Synthetic.
OS Saccharomyces sp.
PN WO9801573-A1.
PD 15-JAN-1998.
PF 09-JUL-1996; U11478.
PR 09-JUL-1996; WO-011478.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI Kourilna NY, Lariouov VL, Perkins EL, Resnick MA;
PI WPI: 98-110234/10.
PT Preparation of yeast artificial chromosomes - by in vivo
PT recombination using vector comprising yeast centromere, marker,
PT repeat telomere and nucleic acid for recombination
PS Example 1; Page 45: 117pp; English.
CC This is the nucleotide sequence for the PCR primer used in the
CC amplification of the Alu repeat sequence, which is used to
CC demonstrate the processes described in the invention. It involves
CC the creation and use of circular yeast artificial chromosome (YAC)
CC to selectively clone specific nucleic acids from a background of
CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
CC They can be used to rapidly isolate human DNA where only a part of the
CC sequence of DNA is known. Using the methods large fragments of DNA can
CC be easily cloned and analysed.
SQ Sequence 40 BP; 7 A; 12 C; 13 G; 8 T;
```

```
Query Match 0.1%; Score 27; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 22242 CCTGTAATCCAGCACTTGGAGGCC 22268
DB 12 CCTGTAATCCAGCACTTGGAGGCC 38
```

```
RESULT 9
ID T23178
ID T23178 standard; cDNA to mRNA; 61 BP.
AC T23178:
DE 30-AUG-1996 (first entry)
DE Human gene signature HUMGS04920.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
```

OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MARS/) MATSUBARA K.
PI (OKUBO/) OKUBO K.
PT Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function. By preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1305; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 61 Bp; 23 A; 12 C; 14 G; 12 T;

Query Match 0.1%; Score 27; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18097 GATCAGCGCACTGCATTCAGCCTGGG 18123
DB 1 GATCAGCGCACTGCATTCAGCCTGGG 27
|||||

RESULT 10
V37197/c
ID V37197 standard; DNA: 80 Bp.
AC V37197.
DN 14-SEP-1998 (first entry)
DE Oligonucleotide sequence of the specification.
KW Genotype; phenotype; molecular evolutionary engineering;
KW functional biopolymer; virus; ss.
OS Synthetic.
PN W09816636-A1.
PD 23-APR-1998.
PF 17-OCT-1997; J03766.
PR 17-OCT-1996; JP-274855.
PA (MITU) MITSUBISHI CHEM CORP.
PI Fusimi Y, Miyamoto E, Nemoto N, Yanagawa H;
DR WPI: 96-261039/23.
PT Virus containing nucleic acid and protein sections - for use in
PT modification and creation of functional bio:polymers such as
PT enzymes, antibodies and ribozymes.
PS Example 1; Page 44; 68pp; Japanese.
CC The present sequence is used in the course of the invention. The
CC specification describes a molecule for bringing together genotype with
CC phenotype (in vitro virus). The molecule contains a nucleic acid
CC fragment having a base sequence corresponding to a genotype, covalently
CC bonded to a protein fragment containing a protein participating in
CC phenotype expression, the 3'-end of the nucleic acid part being bonded
CC to the C-terminus of the protein part via a purmycin moiety. The nucleic
CC acid fragment preferably consists of RNA corresponding to the gene (free
CC from a termination codon), a spacer (such as two-stranded DNA), a peptide
CC adapter (such as a DNA-tRNA hybrid containing an anticodon corresponding
CC to the terminator codon of the gene) and a puromycin-containing cap
CC capable of binding to an amino acid residue. Translation of the virus is

CC performed in a non-cellular (preferably ribosomal) system (e.g. E. coli
CC ribosome), the protein synthesised by the translation attaching to the
CC puromycin cap, resulting in the complete in vitro virus structure. The
CC method is used in molecular evolutionary engineering to optimise function
CC of a functional biopolymer such as an enzyme, antibody or ribozyme, or
CC to generate new functionality. The virus can be used for the functional
CC optimisation of nucleic acid or protein sequences.
SQ Sequence 80 Bp; 7 A; 10 C; 1 G; 62 T;

Query Match 0.1%; Score 27; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29693 ATTAATAAAAAAAAAAAAAAAAAA 29719
DB 66 ATTAATAAAAAAAAAAAAAAAAAA 40
|||||

RESULT 11
V39744
ID V39744 standard; cDNA: 88 Bp.
AC V39744.
DN 28-SEP-1998 (first entry)
DE Microsatellite analysis antisense strand SEQ ID NO:262.
KW Mass spectrometry; diagnosis; detection; biological sample; infection;
KW genetic disease; chromosomal abnormality; identification; heredity;
KW pathogenic organism; telomerase activity; oncogene mutation;
KW cancer-specific sequence; primer; ss.
OS Synthetic.
PN W09820166-A2.
PD 14-MAY-1998.
PF 06-NOV-1997; U20444.
PR 08-OCT-1997; US-947801.
PR 06-NOV-1996; US-744481.
PR 06-NOV-1996; US-744590.
PR 06-NOV-1996; US-746036.
PR 06-NOV-1996; US-746035.
PR 23-JAN-1997; US-786988.
PR 23-JAN-1997; US-787639.
PR 19-SEP-1997; US-933792.
PA (SEOU-) SEQUENOM INC.
PI Braun A, Damhofer-Demar B, Fu D, Higgins GS, Jurinke C,
PI Koster H, Little DP, Lough DW, Siebert CW, Tang X,
PI Van Den Boom D, Xiang G;
DR WPI: 98-286975/25.
PT Sequencing nucleic acid by mass spectrometric analysis - for
PT detecting nucleic acids, telomerase activity, oncogene mutations, or
PT cancer-specific sequences, for diagnosis of disease
PS Example 11; Page 318; 478pp; English.
CC A process has been developed for determining the sequence of a target
CC nucleic acid. The process comprises: (i) generating at least two
CC fragments (F) from the target nucleic acid; and (ii) analysing F by
CC mass spectrometry (MS). The sequences in V39592 are
CC specifically claimed primers for use in the mass spectrometric analysis
CC of the above process. The process is used to detect genetic diseases
CC (e.g. haemophilia, thalassemia, Duchenne muscular dystrophy, Alzheimer's
CC disease, cystic fibrosis and many others) or chromosomal abnormalities
CC (or predisposition); infections and cancers; also for establishing
CC identity and heredity. Particular applications are diagnosis of
CC neuroblastoma, detecting telomerase, determining family relationships
CC and HLA compatibility, and in genetic fingerprinting. Compared with
CC known methods using MS, this process requires fewer specific reagents
CC and is better suited to automation. Extended primers are shorter;
CC primer annealing is more efficient and the process allows detection of
CC many sequences simultaneously. The present sequence represents an
CC oligonucleotide used in an example from the present invention.
SQ Sequence 88 Bp; 56 A; 5 C; 7 G; 20 T;

Query Match 0.1%; Score 27; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23274 AAAAATAAATAAATAAATAA 23300
DB 15 AAAAATAAATAAATAAATAA 41

RESULT 12
T88439/c
ID T88439 standard; DNA: 92 BP.
AC T88439;
DT 12-MAY-1998 (first entry)
DE Padlock probe alphaV from WO9741254.
KW cleavable padlock probe; target nucleic acid; detection;
KM circularised structure; hybridisation; Factor V; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT modified_base 67 /tag= a /note= "attached to an amino group"
FT FT
PD WO9741254-A1.
PN 06-NOV-1997.
PR 30-APR-1997; SE0737.
PA 30-APR-1996; SE-001676.
PI (LAND/) LANDEGREN U.
DR Landegren U;
WPI: 97-549751/50.
PT Detecting target nucleic acid sequence - using probe capable of
PT forming circularised structure, which can be cleaved to form
PT detectable function
PS Example 2; Fig 7; 20pp; English.
CC The present invention describes a method for detecting a target nucleic
CC acid sequence (I) in a sample. The method comprises: (a) hybridising (I)
CC to the probe ends of a probe having 2 free nucleic acid ends, which
CC are at least partially complementary to and capable of hybridising to
CC at least 2 neighbouring regions of (I); (b) covalently connecting the
CC ends of the hybridised probe with each other to form a circularised
CC structure; (c) washing under denaturing conditions, characterised in
CC that the probe is provided with a cleavable or dissociable detectable
CC function; (d) cleaving or dissociating the detectable function; (e)
CC separating probes with connected ends from probes with non-connected
CC ends by washing under denaturing conditions; and (f) detecting the
CC presence and, if desired, location of the remaining probe as indicative
CC of the presence of (I). The present sequence represents a padlock probe,
CC alpha FV, used in an example of the present invention. The method
CC can be used to detect, quantify and distinguish between sequence
CC variants with regard to 1 or several (I) in a sample. It can be used to
CC distinguish between normal and mutated sequence variants associated with
CC disease, for genetic linkage analysis of biallelic markers and to
CC quantify gene expression in a tissue sample. The circularisable probe is
CC designed so that it reports the presence of (I) by allowing a detectable
CC moiety to remain bound only if the probe has been cyclised in a target
CC dependent linking reaction. It also provides for high specificity and
CC background reduction.
CC Sequence 92 BP; 9 A; 9 C; 8 G; 66 T;
SQ

Query Match 0.1%; Score 27; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23696 AAAAAAAAAAAAAAAAAAGAA 29722
DB 69 AAAAAAAAAAAAAAAAAAGAA 43

RESULT 13
ID 075099/c
AC 075099 standard; cDNA: 95 BP.
DT 11-ANG-1995 (first entry)
DE Plasmid pOKSCLba confg. portion of psychosis protecting protein gene.
KM Psychosis protecting protein; PCR; primer; ss.
OS Homo sapiens.

FH Key Location/Qualifiers
FT misc_difference 1. .28 /tag= a
FT FT /note= "PCR primer used to generate cDNA library"
FT misc_difference 74. .95 /tag= b
FT FT /note= "see above"
FT FT
PN WO9426107-A.
PD 24-NOV-1994.
PE 13-MAY-1994. 005445.
PR 13-MAY-1993; US-060560.
PA (UYNV) UNIV NEW YORK STATE.
PI Basham DA, Friedhoff AJ, Miller JC;
DR WPI: 95-006234/01
PT New nucleic acids encoding psychosis protecting peptide and
PT antibodies - for the treatment, diagnosis and research of
PT psychotic disorders, such as schizophrenia
PS Claim 1; Page 52; 87 pp; English.
CC A subtracted cDNA library was generated from the RNA of lymphocytes
CC obtd. from monozygotic twins discordant for schizophrenia. The two
CC oligos used in library construction were T7 promoter and SP6
CC promoter (see 075091, 075092). A cDNA library was made for each
CC twin and subtractive hybridisation was achieved. Both libraries
CC were used as driver and substrate in two separate subtraction
CC assays. Clone pOKS18a was isolated and sequenced.
SQ Sequence 95 BP; 12 A; 14 C; 30 G; 39 T;

Query Match 0.1%; Score 27; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24134 GAAGTCGATCTCAAAAAAAAAA 24160
DB 41 GAAGTCGATCTCAAAAAAAAAA 15

RESULT 14
ID V68827/c
AC V68827 standard; DNA: 100 BP.
DT 22-JAN-1999 (first entry)
DE DNA molecule encoding a breast tumour specific polypeptide #19.
KM Human; breast cancer; breast tumour tissue; diagnosis; treatment;
KM vaccine; epitope; endogenous; retroviral element; ss.
OS Homo sapiens.
PN WO9845328-A2.
PD 15-OCT-1998.
PE 09-APR-1998; 006939.
PR 11-DEC-1997; US-991789.
PR 09-APR-1997; US-838762.
PA (CORI-) CORIXA CORP.
PI Frudakis TN, Reed SG, Smith JM;
DR WPI: 98-557473/47.
PT New DNA sequences isolated from endogenous human retroviral element
PT - and related vectors, transformed cells, proteins and antibodies,
PT useful for diagnosis, treatment and prevention of breast cancer
PS Claim 1; Page 47; 173pp; English.
CC V68800 to V68998 represent nucleotide sequences which encode human
CC breast tumour specific polypeptides. Detection or measurement of
CC human breast tumour specific polypeptides and nucleotide sequences,
CC or the corresponding RNA in a sample, is used for diagnosis and
CC monitoring of breast cancer. Human breast tumour specific polypeptides
CC and nucleotide sequences, and the vectors containing the DNAs, are also
CC useful in vaccines for inhibiting development (for prevention or
CC therapy) of breast cancer. The polypeptides may also be used to
CC raise monoclonal antibodies, used as immunoassay reagents.
SQ Sequence 100 BP; 19 A; 10 C; 8 G; 63 T;

Query Match 0.1%; Score 27; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Mon Apr 3 08:24:03 2000

us-08-852-495c-1_copy_115000_145000.rng

Page 6

OY 29696 AAAAAAAAAAAAAAAAAAGAA 29722
DB 74 AAAAAAAAAAAAAAAAAAGAA 48

RESULT 15

V71935/C
ID V71935 standard; DNA; 27 BP.
AC V71935.
DE 18-FEB-1999 (first entry)
DE Anchored poly (RT-PCR primer).
KW Normalised; cDNA library; mRNA cloning; reverse transcription;
KW Immobilised; screening; hybridisation; nucleic acid amplification;
KW expression pattern; drug development; PCR primer; RT-PCR; ss.
OS Synthetic.
PN W09851789-A2.
PD 19-NOV-1998.
PE 13-MAY-1998; DK0186.
PR 27-MAR-1998; DK-000432.
PR 13-MAY-1997; DK-000547.
PR 19-MAY-1997; US-871030.
PA (DISP-) DISPLAY SYSTEMS BIOTECH APS.
PI Warthoe PR:
DR WPI: 99-009772/01.
PR Preparation of normalised, subdivided cDNA libraries from mRNA - by
PR reverse transcription and amplification, used to screen for new
PR genes and interacting proteins, potential drugs, and for diagnosis
PR Example 1; Page 29; 71pp; English.
CC The invention relates to preparation of a normalised, subdivided library
CC of amplified cDNA from the coding regions of mRNA in a sample. The method
CC involves reverse transcription with at least one cDNA primer of formula
CC of 10n1-dn2-Vn3-Nn4 to form first strand cDNA where Con1 - any sequence
CC are 100 nucleotides; dn - deoxythymidyl, n2 is at least 1; n3 and n4
CC cDNA both 0, or n3 is 1 and n4 is at least 1; followed by second strand
CC primer synthesis using the first strand as template and a second cDNA
CC primer of a similar formula, in the presence of DNA polymerase I (or its
CC amplification fragment) and amplification of double stranded cDNA with a set of
CC database computer-generated list of molecular weights of restricted
CC DNA fragments of known sequence) is used to determine presence of an
CC expressed protein in a cell, also to detect changes in such expression.
CC (particularly for diagnosis of disease). Surfaces having amplified
CC cDNA stably immobilised on it, obtained by a similar method, are used
CC to screen for genes of a particular family, by hybridisation with nucleic
CC acid from the family (to identify new genes) and to detect differences in
CC expression patterns between cells. The polypeptides expressed by the
CC libraries can be used for drug development. Sequences V71935 to V71946
CC represent primers used to exemplify the method of the invention.
SQ Sequence 27 BP; 2 A; 0 C; 0 G; 25 T;

Query Match 0.18; Score 26; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 29694 TTTAAAAAAAAAAAAAAAAAAAAA 29719
DB 27 TTTAAAAAAAAAAAAAAAAAAAAA 2

Search completed: March 31, 2000, 00:01:30
Job time: 159088 sec

Mon Apr 3 08:24:04 2000

us-08-852-495c-1_copy_115000_145000.rni

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: March 30, 2000, 17:28:35 ; Search time 345.24 Seconds

(without alignments)
10403.818 Million cell updates/sec

Title: US-08-852-495C-1_COPY_115000_145000

Perfect score: 30001
Sequence: 1 AATCTTCACATTTTGATA.....AGAAATACCTATTAA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 214294 seqs, 59861574 residues

Word size : 0

Total number of hits satisfying chosen parameters: 349160

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/lna/5C.COMB.seq:*
- 4: /cgn2_6/ptodata/1/lna/5D.COMB.seq:*
- 5: /cgn2_6/ptodata/1/lna/6.COMB.seq:*
- 6: /cgn2_6/ptodata/1/lna/6.PCT059.COMB.seq:*
- 7: /cgn2_6/ptodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	38	0.1	60	3	US-08-454-557C-57	Sequence 57, Appl
2	38	0.1	60	4	US-08-340-426D-57	Sequence 57, Appl
3	38	0.1	60	4	US-08-450-673C-57	Sequence 57, Appl
4	38	0.1	60	6	PCR-US95-17111A-57	Sequence 57, Appl
5	38	0.1	76	3	US-08-454-557C-69	Sequence 69, Appl
6	38	0.1	76	4	US-08-340-426D-69	Sequence 69, Appl
7	38	0.1	76	4	US-08-450-673C-69	Sequence 69, Appl
8	38	0.1	76	6	PCR-US95-17111A-69	Sequence 69, Appl
9	28	0.1	50	1	US-08-233-609-5	Sequence 5, Appl
10	28	0.1	50	1	US-08-444-083-5	Sequence 5, Appl
11	28	0.1	50	1	US-08-286-304-5	Sequence 5, Appl
12	28	0.1	50	1	US-08-442-745-5	Sequence 5, Appl
13	28	0.1	50	1	US-08-443-129-5	Sequence 5, Appl
14	28	0.1	50	1	US-08-443-952-5	Sequence 5, Appl
15	28	0.1	50	2	US-08-443-130-5	Sequence 5, Appl
16	28	0.1	50	6	PCR-US95-04467-5	Sequence 5, Appl
17	28	0.1	57	3	US-08-332-766A-3	Sequence 3, Appl
18	26	0.1	35	1	US-08-253-889-10	Sequence 10, Appl
19	26	0.1	42	1	US-07-875-167-2	Sequence 2, Appl
20	26	0.1	42	1	US-08-287-164-2	Sequence 2, Appl
21	26	0.1	48	2	US-08-741-881-21	Sequence 21, Appl
22	26	0.1	48	2	US-08-739-158-21	Sequence 21, Appl
23	26	0.1	48	3	US-08-739-167-21	Sequence 21, Appl
24	26	0.1	92	1	US-08-120-827-94	Sequence 94, Appl
25	26	0.1	92	2	US-08-478-675-94	Sequence 94, Appl
26	25	0.1	26	2	US-08-621-914A-1	Sequence 1, Appl

27	25	0.1	33	7	5478746-1	Patent No. 5478746
28	25	0.1	41	1	US-08-113-646A-39	Sequence 39, Appl
29	25	0.1	44	1	US-08-113-646A-40	Sequence 40, Appl
30	25	0.1	46	1	US-08-222-177A-349	Sequence 349, App
31	25	0.1	47	4	US-08-778-494B-114	Sequence 114, App
32	25	0.1	48	2	US-08-380-438-6	Sequence 6, Appl
33	25	0.1	50	1	US-08-381-572-20	Sequence 20, Appl
34	25	0.1	50	2	US-08-593-820-20	Sequence 20, Appl
35	25	0.1	50	4	US-08-663-823B-72	Sequence 72, Appl
36	25	0.1	51	4	US-08-582-562A-8	Sequence 8, Appl
37	25	0.1	51	4	US-08-778-494B-8	Sequence 8, Appl
38	25	0.1	51	4	US-08-859-998-1373	Sequence 1373, App
39	25	0.1	52	4	US-08-778-494B-111	Sequence 111, App
40	25	0.1	54	4	US-08-771-624B-24	Sequence 24, Appl
41	25	0.1	55	1	US-08-113-646A-41	Sequence 41, Appl
42	25	0.1	55	4	US-08-522-269B-18	Sequence 18, Appl
43	25	0.1	55	4	US-08-582-562A-16	Sequence 16, Appl
44	25	0.1	55	4	US-08-778-494B-16	Sequence 16, Appl
45	25	0.1	55	4	US-08-778-494B-67	Sequence 67, Appl

ALIGNMENTS

RESULT 1
US-08-454-557C-57
Sequence 57, Application US/08454557C
Patent No. 5630670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-57
Query Match 0.1%; Score 38; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.00064;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17232 TTTTGTATTTTGTAGACAGGCTTACCGTCTT 17269
DB 16 TTTTGTATTTTGTAGACAGGCTTACCGTCTT 53

Query Match 0.1%; Score 38; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.00064;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17232 TTTTGTATTTTAGTAGACAGCGGTTTCACCGTGT 17269
|||||
DB 16 TTTTGTATTTTAGTAGACAGCGGTTTCACCGTGT 53

RESULT 5

US-08-454-557C-69
; Sequence 69, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-454-557C-69

Query Match 0.1%; Score 38; DB 3; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.00061;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17232 TTTTGTATTTTAGTAGACAGCGGTTTCACCGTGT 17269
|||||
DB 16 TTTTGTATTTTAGTAGACAGCGGTTTCACCGTGT 53

RESULT 6

US-08-340-426D-69
; Sequence 69, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-340-426D-69

Query Match 0.1%; Score 38; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.00061;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17232 TTTTGTATTTTAGTAGACAGCGGTTTCACCGTGT 17269
|||||
DB 16 TTTTGTATTTTAGTAGACAGCGGTTTCACCGTGT 53

RESULT 7

US-08-450-673C-69
; Sequence 69, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540

Mon Apr 3 08:24:04 2000

us-08-852-495c-1_copy_115000_145000.rn1

Page 4

INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-69

Query Match
Best Local Similarity 100.0%; Score 38; DB 4; Length 76;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17232 TTTTGTATTTTGTAGACAGAGGTTTCACCGTGT 17269
DB 16 TTTTGTATTTTGTAGACAGAGGTTTCACCGTGT 53

RESULT 8
PCT-US95-17111A-69
Sequence 69, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
ZIP: 20005-3934
COUNTRY: U.S.A.
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609,3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-69

Query Match
Best Local Similarity 100.0%; Score 38; DB 6; Length 76;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17232 TTTTGTATTTTGTAGACAGAGGTTTCACCGTGT 17269
DB 16 TTTTGTATTTTGTAGACAGAGGTTTCACCGTGT 53

RESULT 9
US-08-233-609-5/C

Sequence 5, Application US/08233609
Patent No. 5534615
GENERAL INFORMATION:
APPLICANT: Baker, Joffre
APPLICANT: Chien, Kenneth
APPLICANT: King, Kathleen
APPLICANT: Pennica, Diane
APPLICANT: Wood, William
TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/233,609
FILING DATE: 25-APR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 894
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-233-609-5

Query Match
Best Local Similarity 100.0%; Score 28; DB 1; Length 50;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29696 AAAAAAAAAAAAAAAAAAAAAAGANT 29723
DB 44 AAAAAAAAAAAAAAAAAAAAAAGANT 17

RESULT 10
US-08-444-083-5/C
Sequence 5, Application US/08444083
Patent No. 5571675
GENERAL INFORMATION:
APPLICANT: Baker, Joffre
APPLICANT: Chien, Kenneth
APPLICANT: King, Kathleen
APPLICANT: Pennica, Diane
APPLICANT: Wood, William
TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA

ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,083
FILING DATE: 17-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/233609
FILING DATE: 25-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/286304
FILING DATE: 05-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: 894P1D5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-444-083-5

Query Match 0.1%; Score 28; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29696 AAAAAAAAAAAAAAAAAAGAAAT 29723
DB 44 AAAAAAAAAAAAAAAAAAGAAAT 17

RESULT 11
US-08-286-304-5/C
Sequence 5, Application US/08286304
Patent No. 5571893
GENERAL INFORMATION:
APPLICANT: Baker, Joffre
APPLICANT: Chien, Kenneth
APPLICANT: King, Kathleen
APPLICANT: Pennica, Diane
APPLICANT: Wood, William
TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,304
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/233609
FILING DATE: 25-APR-1994

ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 894P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-286-304-5

Query Match 0.1%; Score 28; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29696 AAAAAAAAAAAAAAAAAAGAAAT 29723
DB 44 AAAAAAAAAAAAAAAAAAGAAAT 17

RESULT 12
US-08-442-745-5/C
Sequence 5, Application US/08442745
Patent No. 5624806
GENERAL INFORMATION:
APPLICANT: Baker, Joffre
APPLICANT: Chien, Kenneth
APPLICANT: King, Kathleen
APPLICANT: Pennica, Diane
APPLICANT: Wood, William
TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,745
FILING DATE: 17-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/233609
FILING DATE: 25-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/286304
FILING DATE: 05-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: 894P1D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 bases
TYPE: nucleic acid
STRANDEDNESS: single

Mon Apr 3 08:24:04 2000

us-08-852-495c-1_copy-115000_145000.rml

TOPOLOGY: linear

Query Match 0.1%; Score 28; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29696 AAAAAAAAAAAAAAAAAAGAAAT 29723
DB 44 AAAAAAAAAAAAAAAAAAGAAAT 17

RESULT 13
US-08-443-129-5/c
Sequence 5, Application US/08443129
Patent No. 5627073
GENERAL INFORMATION:
APPLICANT: Baker, Joffie
APPLICANT: Chien, Kenneth
APPLICANT: King, Kathleen
APPLICANT: Pennica, Diane
APPLICANT: Wood, William
TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/443.129
FILING DATE: 17-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/233609
FILING DATE: 25-APR-1994
PRIOR APPLICATION DATA: 08/286
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: 304
FILING DATE: (null)
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36700
REFERENCE/DOCKET NUMBER: 894P1D2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-443-129-5

Query Match 0.1%; Score 28; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29696 AAAAAAAAAAAAAAAAAAGAAAT 29723
DB 44 AAAAAAAAAAAAAAAAAAGAAAT 17

DB 44 AAAAAAAAAAAAAAAAAAGAAAT 17

RESULT 14
US-08-443-952-5/c
Sequence 5, Application US/08443952
Patent No. 5679545
GENERAL INFORMATION:
APPLICANT: Baker, Joffie
APPLICANT: Chien, Kenneth
APPLICANT: King, Kathleen
APPLICANT: Pennica, Diane
APPLICANT: Wood, William
TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses Therefor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/443.952
FILING DATE: 17-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/233609
FILING DATE: 25-APR-1994
PRIOR APPLICATION DATA: 08/286304
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: 304
FILING DATE: (null)
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Timothy E.
REGISTRATION NUMBER: 36700
REFERENCE/DOCKET NUMBER: 894P1D4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-443-952-5

Query Match 0.1%; Score 28; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29696 AAAAAAAAAAAAAAAAAAGAAAT 29723
DB 44 AAAAAAAAAAAAAAAAAAGAAAT 17

RESULT 15
US-08-443-130-5/c
Sequence 5, Application US/08443130
Patent No. 5723585
GENERAL INFORMATION:
APPLICANT: Baker, Joffie
APPLICANT: Chien, Kenneth
APPLICANT: King, Kathleen
APPLICANT: Pennica, Diane
APPLICANT: Wood, William

US-08-443-130-5

```

: TITLE OF INVENTION: Cardiac Hypertrophy Factor and Uses
: TITLE OF INVENTION: Therefor
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 Inch, 360 KB floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/443,130
: FILING DATE: 17-MAY-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/233609
: FILING DATE: 25-APR-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/286304
: FILING DATE: 05-AUG-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Torchia, Timothy E.
: REGISTRATION NUMBER: 36,700
: REFERENCE/DOCKET NUMBER: 894PID3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-8674
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 50 Bases
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: US-08-443-130-5

```

```

Query Match          0.1%; Score 28; D3 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 29696 AAAAAAAAAAAAAAAAAAGAAAT 29723
      |||||||||||||||||||||||||
Db    44 AAAAAAAAAAAAAAAAAAGAAAT 17

```

```

Search completed: March 30, 2000, 23:51:03
Job time: 158525 sec

```



1
2
3

4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 13:20:07 ; Search time 8198.72 Seconds
(without alignments)
13816.022 Million cell updates/sec

Title: US-08-852-495c-1_COPY_115000_145000
Perfect score: 30001
Sequence: 1 AAATCTTCGATTGTGATA.....AGAAATAACCGATTAA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : EST:*

1:	em_est1:*
2:	em_est2:*
3:	em_est3:*
4:	em_est4:*
5:	em_est5:*
6:	em_est6:*
7:	em_est7:*
8:	em_est8:*
9:	em_est9:*
10:	em_est10:*
11:	em_est11:*
12:	em_est12:*
13:	em_est13:*
14:	em_est14:*
15:	em_est15:*
16:	em_est16:*
17:	em_est17:*
18:	em_est18:*
19:	em_est19:*
20:	gb_est1:*
21:	gb_est2:*
22:	gb_est3:*
23:	gb_est4:*
24:	gb_est5:*
25:	gb_est6:*
26:	gb_est7:*
27:	gb_est8:*
28:	gb_est9:*
29:	gb_est10:*
30:	gb_est11:*
31:	gb_est12:*
32:	gb_est13:*
33:	gb_est14:*
34:	gb_est15:*
35:	gb_est16:*
36:	gb_est17:*
37:	gb_est18:*
38:	gb_est19:*
39:	gb_est20:*
40:	gb_est21:*
41:	gb_est22:*
42:	gb_est23:*
43:	gb_est24:*

44:	gb_est25:*
45:	gb_est26:*
46:	gb_est27:*
47:	gb_est28:*
48:	gb_est29:*
49:	gb_est30:*
50:	gb_est31:*
51:	gb_est32:*
52:	em_est20:*
53:	em_est21:*
54:	em_est22:*
55:	em_est23:*
56:	em_est24:*
57:	em_est25:*
58:	em_est26:*
59:	gb_est33:*
60:	gb_est34:*
61:	gb_est35:*
62:	gb_est36:*
63:	gb_est37:*
64:	gb_est38:*
65:	em_est27:*
66:	em_est28:*
67:	em_est29:*
68:	em_est30:*
69:	gb_est39:*
70:	gb_est40:*
71:	gb_est41:*
72:	gb_est42:*
73:	gb_est43:*
74:	gb_est44:*
75:	em_est31:*
76:	em_est32:*
77:	em_est33:*
78:	em_est34:*
79:	gb_est45:*
80:	gb_est46:*
81:	gb_est47:*
82:	gb_est48:*
83:	em_est35:*
84:	em_est36:*
85:	em_est37:*
86:	em_est38:*
87:	gb_est49:*
88:	gb_est50:*
89:	gb_est51:*
90:	gb_est52:*
91:	gb_est53:*
92:	em_est39:*
93:	em_est40:*
94:	em_est41:*
95:	em_est42:*
96:	em_est43:*
97:	em_est44:*
98:	em_est45:*
99:	gb_est54:*
100:	gb_est55:*
101:	em_est46:*
102:	gb_est56:*
103:	gb_est57:*
104:	gb_est58:*
105:	gb_est59:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	44	0.1	101	100	A0280224
					A0280224 CITBI-EI-

C 2 42 0.1 93 43 A1168167
 C 3 41 0.1 95 20 T52775
 C 4 39 0.1 97 39 AA837701
 C 5 38 0.1 102 100 A0319270
 C 6 38 0.1 102 36 AA654562
 C 7 38 0.1 103 104 A0584425
 C 8 37 0.1 82 34 AA58985
 C 9 37 0.1 85 50 F26823
 C 10 37 0.1 103 35 AA570476
 C 11 37 0.1 103 81 B90619
 C 12 37 0.1 102 104 A0535244
 C 13 36 0.1 97 20 A0535244
 C 14 36 0.1 97 20 A0535244
 C 15 35 0.1 65 81 B36140
 C 16 35 0.1 101 100 A0278892
 C 17 35 0.1 105 100 A0278892
 C 18 34 0.1 65 25 AA65303
 C 19 34 0.1 65 25 AA65303
 C 20 34 0.1 101 33 AA581369
 C 21 34 0.1 104 29 AA128957
 C 22 33 0.1 91 38 AA670754
 C 23 33 0.1 95 37 AA670754
 C 24 33 0.1 101 21 R1703571
 C 25 33 0.1 101 25 N80946
 C 26 33 0.1 101 25 N80946
 C 27 33 0.1 102 81 B48088
 C 28 33 0.1 103 104 A0541035
 C 29 33 0.1 105 28 A078003
 C 30 32 0.1 71 31 AA81256
 C 31 32 0.1 86 36 AA65098
 C 32 32 0.1 102 81 B62983
 C 33 32 0.1 102 81 B62983
 C 34 32 0.1 104 50 A1675576
 C 35 32 0.1 52 20 D25845
 C 36 31 0.1 53 43 A1206968
 C 37 31 0.1 84 34 AA501753
 C 38 31 0.1 95 30 T52775
 C 39 31 0.1 100 35 AA583697
 C 40 31 0.1 101 39 AA835205
 C 41 31 0.1 103 74 N39933
 C 42 31 0.1 103 74 N39933
 C 43 31 0.1 103 74 N39933
 C 44 31 0.1 103 74 N39933
 C 45 31 0.1 103 81 B48914

ALIGNMENTS

A1168167 o09e10.x
 T52775 ya79408.r1
 AA837701 o06c02.s
 A0319270 RPCI11-98
 AA654562 nt75f10.s
 A0584425 RPCI11-4
 AA58985 zx8108.s
 F26823 HSPD1446.H
 AA570476 nk6312.s
 B90619 CIT-HSP-216
 A0535244 RPCI11-3
 D0989 HUMS01971
 A1675576 w02c03.x
 B36140 HS-1038-A1-
 A0278892 CITBI-E1-
 A0278892 RPCI11-94
 AA65303 CG88a05.s
 AA128957 EST94442
 AA670754 zns6h04.r
 AA670754 ac8812.s
 AA659571 AC8810.s
 R17033 Y4C84D10.s
 N80946 Y412506.r2
 N80946 Y412506.r2
 B48088 RPCI11-4N6
 A0541035 RPCI11-3
 A078003 7H12D08.C
 H92864 Y19C11.S1
 AA281256 z894e01.s
 AA65098 nt76d11.s
 B62983 CIT-HSP-669
 AA807640 nx08b05.s
 A1675576 w02c03.x
 D25845 HUMS04217
 A1206968 q10g11.x
 AA501753 ng05a01.s
 AA291270 z818d04.s
 T52775 ya79408.r1
 AA583697 n122a06.s
 AA564832 nt58f10.s
 AA835205 ak64h01.s
 N39933 Y68911.r1
 AA176421 QVO-C7001
 B48914 RPCI11-4A12

RESULT 1
 LOCUS A0280224/c
 DEFINITION A0280224 101 bp DNA
 CITBI-E1-2522N7.TF CITBI-E1 Homo sapiens genomic clone 2522N7,
 genomic survey sequence.
 VERSION A0280224.1 GI:3906043
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 101)
 Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
 Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
 Venter,J.C.
 Title Use of a random human BAC End Sequence Database for Sequence-Ready
 Map Building (1998)
 Other-GSS: CITBI-E1-2522N7.TF
 Contact: Mark Adams
 The Institute of Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA

FEATURES

source
 1..101
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1lb="CITBI-E1"
 /sex="male"
 /cell_type="sperm"
 /note="Vector: pReloBAC11, Site:1, EcoRI; Site:2, EcoRI;
 Caltech Human BAC Library D"
 Class: BAC ends

BASE COUNT 24 a 28 c 24 g 25 t
 ORIGIN

Query Match 0.1%; Score 44; DB 100; Length 101;
 Best Local Similarity 100.0%; Pred No. 0.00059;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 22230 CAGGGCTCACACCTGTATCCAGCAGCTTGGAGCCAGAGT 22273
 DB 70 CAGTGGCTCACACCTGTATCCAGCAGCTTGGAGCCAGAGT 27

RESULT 2
 LOCUS A1168167/c
 DEFINITION A1168167 93 bp mRNA
 IMAGE:1555706.3 similar to contains Alu repetitive
 element, contains TARI repetitive element, mRNA sequence.
 VERSION A1168167.1 GI:3701337
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 93)
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 On Jan 19, 1998 this sequence version replaced gi:2282306.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 1481 bp
 Seq primer: -40m13 fwd: ET, rev: from Amersham
 High quality sequence stop 75.
 Location/Qualifiers
 1..93
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1lb="IMAGE:1555706"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: p773D-Pac (Pharmacia) with
 Equal amounts of plasmid DNA from five normalized
 libraries were mixed and ss circles were made in vitro.
 a subtractive hybridization, this DNA was used as tracer in
 PCR-amplified cDNAs from pools of 5,000 clones made from
 the same 5 libraries. The pools consisted of the following

FEATURES
 source


```
DB      97 GAGATCACTTGACCTGGGAGGACGAGTTGTCATGAG 59

RESULT  5
LOCUS   AQ319270      102 bp      DNA      GSS      06-MAY-1999
DEFINITION RPc11-98B22.TJ RPc1-11 Homo sapiens genomic clone RPc1-11-98B22,
            genomic survey sequence.
ACCESSION AQ319270
VERSION   AQ319270.1 GI:4052235
KEYWORDS GSS.
SOURCE   human.
ORGANISM Homo sapiens
            Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 102)
AUTHORS   Adams M.D., Rounsley S.D., Zhao S., Bass S., Linher K., Golden K.,
            Berry K., Granger D., Suh E., Mible C., de Jong P. and Venter J.C.
            Use of human BAC End sequences for Sequence-Ready Map Building
            Unpublished (1999)
TITLE     Other-GSS: RPc11-98B22.TJ
JOURNAL   Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPc1-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@ed.jong.med.bufileo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.bufileo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.
FEATURES
            source
            Location/Qualifiers
                1..102
                /organism="Homo sapiens"
                /db_xref="GDB:7537293"
                /db_xref="taxon:9606"
                /clone="RPc1-11-98B22"
                /clone_11b="RPc1-11"
                /sex="Male"
                /cell-type="Lymphocytes"
                /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
                RPc111 Human Male BAC Library"
BASE COUNT      26 a      28 c      31 g      17 t
ORIGIN
Query Match      0.1%; Score 39; DB:100; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22230 CAGTGGCTACACCTGTATCCGACACTTGGAGGCC 22268
|||||
DB      42 CAGTGGCTACACCTGTATCCGACACTTGGAGGCC 80

RESULT  6
LOCUS   AA654562      102 bp      mRNA      EST      04-NOV-1997
DEFINITION nt75110.s1 NCI_CGAP_P33 Homo sapiens cDNA clone IMAGE:1204363
            similar to contains Alu repetitive element;contains element MER22
            repetitive element ; , mRNA sequence.
ACCESSION AA654562
VERSION   AA654562.1 GI:2590716
KEYWORDS EST.
SOURCE   human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 102)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
            On Sep 12, 1996 this sequence version replaced gi:1393453.
COMMENT    Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert.Strausberg@nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquag,
            M.D., Michael Emmert-Buck, M.D., Ph.D.
            CDNA Library Preparation: David B. Kitzman, Ph.D.
            CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www.bio.llnl.gov/bbrp/image/image.html
FEATURES
            source
            Location/Qualifiers
                1..102
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_11b="NCI_CGAP_P33"
                /clone="IMAGE:1204363"
                /sex="Male"
                /dev_stage="45 years old"
                /lab_host="DH10B"
                /note="Vector: PAMP10; Site_1: NotI; Site_2: EcoRI; 1st
                strand cDNA was primed with oligo(dT)17 on 50 ng of
                DNase-treated, total cellular RNA obtained from
                5,000-10,000 microdissected cells
                histologically determined to be fully malignant prostate
                cancer cells. Double-stranded cDNA was ligated to EcoRI
                adaptors, 5 cycles of PCR applied to the cDNA with an
                adaptor-specific primer, and the resulting PCR product
                subcloned into PAMP10 by the UBE-cloning method (Life
                Technologies). Average insert size is 600 bp. NOTE: Not
                directionally cloned. This library was constructed by
                David Kitzman."
BASE COUNT      22 a      32 c      27 g      21 t
ORIGIN
Query Match      0.1%; Score 38; DB:36; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10130 GGAAGTTGAGACCAAGCTGGCCACATGTGAACCTC 10167
|||||
DB      42 GGAAGTTGAGACCAAGCTGGCCACATGTGAACCTC 5

RESULT  7
LOCUS   A0584425      103 bp      DNA      GSS      07-JUN-1999
DEFINITION RPc1-11-45812.TJ RPc1-11 Homo sapiens genomic clone RPc1-11-45812,
            genomic survey sequence.
ACCESSION A0584425
VERSION   A0584425.1 GI:5011535
KEYWORDS GSS.
SOURCE   human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS   Zhao S., Adams M.D., Nierman W., Malek J., de Jong P. and
            Venter J.C.
            Use of BAC End sequences from library RPc1-11 for Sequence-Ready
            Map Building
            Unpublished (1997)
            Other-GSS: RPc1-11-45812.TJ
            Contact: Shaying Zhao, William Nierman, Mark Adams
JOURNAL   CONTACT: Shaying Zhao, William Nierman, Mark Adams
```

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208

Clones are derived from the human BAC library RCI1-11. For BAC library availability, please contact pierer de jong (pierredejong.med.buffalo.edu). Clones may be purchased from BACpac Resources (<http://Bacpac.med.buffalo.edu/ordering>) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_and_search/bac_end_search.html

Seq primer: Sp6

Class: BAC ends.

FEATURES

```

source      1. .103
            /organism="Homo sapiens"
            /db_xref="GDB:767573"
            /db_xref="taxon:9606"
            /clone="RPC1-11.458L2"
            /clone_1lb="RPC1-11"
            /sex="Male"
            /cell_type="Lymphocytes"
            /note="Vector: pBAC3.6; Site.1: EcoRI; Site.2: EcoRI;
            RPc11 Human Male BAC Library"
BASE COUNT 23 a      27 c      31 g      22 t
ORIGIN

```

Query Match	0.18;	Score 38;	DB 104;	Length 103;
Best Local Similarity	100.08;	Pred. NO. 0.05;		
Matches 38;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 16505 TGATCTGCCCACTTGGCCTCCCAAGTGTGGATAA 16542

Db 56 TGATCTGCCACCTTGGCCTCCCAAAGTCTGGGATA 19

RESULT 8

LOCUS	82 bp	mRNA	EST	09-JUN-1997
DEFINITION	AA458985			
IMAGE:810879	3'	similar to contains LTRs.	t2 LTR5 repetitive element	
;;	mRNA sequence.			

ACCESSION	AA458985
VERSION	AA458985.1
GI	GI:2183892

SOURCE

ORGANISM

REFERENCE
AUTHORS

Euarystota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 82)
Hillier, L., Allen, M., Bowles, L., Dubouque, T., Geisels, G., Jost

Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Scheellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1395423

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: estewatson.wustl.edu
This clone is available royalty-free through LINL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -41m13 fwd. Et from Amersham
High quality sequence stop: 70.

FEATURES

```
source 1. 82
        /organism="Homo sapiens"
        /db_xref="GDB:6041714"
```

```

/db_xref="taxon:9606"
/clone="IMAGE:810879"
/clone_lib="Soares ovary tumor NBHOT"
/sex="Female"
/tissue_type="Ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: ovary; Vector: pRT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTTCCATCTGATGAAGTGAGCGGCGCGGGTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pRT73 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo."
```

```

Qy 19894 CAAGGCAGAGAATTTTCTTAGTCAGAACAAATG 19930
      |||||
Db 9 CAAGGCAGAGAATTTTCTTAGTCAGAACAAATG 45

```

RESULT 9

LOCUS	85 bp	EST	13-MAY-1999
F26823	85 bp	EST	13-MAY-1999
DEFINITION	HSPD14446	HM3 Homo sapiens cDNA clone s4000060610, mRNA sequence	

VERSION F26823.1 GI:4812449

SOURCE human.

ORGANISM

REFERENCE
1 (bases 1 to 85)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia
Eucheria; Primates; Catarrhini; Homnidae; Homo.

TITLE	AUTHORS
Identification of 4370 expressed sequence tags from a	Lanfranchi,G., Muraro,T., Caldera,F., Pacchioni,B., Pallavicini,A.
	Pandolfi,D., Toppo,S., Trevisan,S., Scarzo,S. and Valle,G.

3'-end-specific cDNA library of human sequencing and filter hybridization
Genome Res. 6 (1), 35-42 (1996)
JOURNAL

COMMENT

CRIISI Biotechnology Centre
University of Padua
VIA TRIESTE 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
<http://grup.bio.unipd.it>

FEATURES	Location/Qualifiers
source	1. .85

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="s400060610"
```

```

/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/notes="Vector: pcDNAIII (Invitrogen); Site_1: BstXI;
Site_2: NotI; The library was constructed by G.

```

Lanfranchi. This library is not subtracted nor normalized. The first strand cDNA was primed with a biotinylated oligo-dT-NotI primer (5'-gttTAACCGGCGGAGCGGCGCTTTTTTTTTTTTTTTT-3'). The ds cDNA was sonicated and size-selected in the range 350-550 bp. The 3' specific fragments were selected by streptavidin coated magnetic beads, ligated to non-palindromic BstXI adapters, NotI digested and directionally cloned into BstXI-NotI cut pCDNAII vector."

BASE COUNT 19 a 18 c 16 g 32 t
ORIGIN

Query Match 0.1%; Score 37; DB 50; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18994 CAGGCGAGAGATTTTCTTAGTACAGACAAATG 19930
|||||
DB 55 CAGGCGAGAGATTTTCTTAGTACAGACAAATG 19

RESULT 10
AA570476 103 bp mRNA EST 09-SEP-1997
LOCUS nk63h12.s1 NCI-CGAP_schl Homo sapiens cDNA clone IMAGE:1018247 3'
DEFINITION Similar to contains Alu repetitive element; mRNA sequence.
ACCESSION AA570476
VERSION AA570476.1 GI:2344456
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index (1997)
JOURNAL Unpublished (1997)
COMMENT On May 8, 1995 this sequence version replaced gi:801194.
CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Stratagene, Inc., David B. Krizman,
Ph.D.
CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILM at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 2104 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amer sham
High quality sequence stop: 76.
Location/Qualifiers

FEATURES
source 1..103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1018247"
/clone_1bp="NCI CGAP_schl"
/issue_type="Schwannoma tumor"
/lab_host="SOLR (Kanamycin resistant)"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. 5' adaptor sequence: 5'
GAATTCGACGAG 3' 3' adaptor sequence: 5'
CTGAGCTTTTCTTTTCTTTT 3' Average insert size: 1.2 kb."
BASE COUNT 27 a 19 c 31 g 26 t
ORIGIN

Query Match 0.1%; Score 37; DB 35; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17877 TTGGCTATGCGTATCCGAGTCTTGGAGGCC 17913
|||||
DB 50 TTGGCTATGCGTATCCGAGTCTTGGAGGCC 86

RESULT 11
B90619 103 bp DNA GSS 25-JUN-1998
LOCUS CTR-HSP-2163G1.TF CTR-HSP Homo sapiens genomic clone 2163G1,
DEFINITION genomic survey sequence.
ACCESSION B90619
VERSION B90619.1 GI:2973099
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE Adams, M.D., Rounsley, S.D., Zhao, S., Field, C.E., Baas, S., Linher, K.,
1 (bases 1 to 103)
Golden, R., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,
Simon, M. and Venter, J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL Unpublished (1998)
COMMENT Other GSSs: CTR-HSP-2163G1.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungem/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES

source 1..103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2163G1"
/clone_1bp="CTR-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT 27 a 31 c 27 g 18 t
ORIGIN

Query Match 0.1%; Score 37; DB 81; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 22232 GTGGCTACACCTGATCCGAGCTTGGAGGCC 22268
|||||
DB 65 GTGGCTACACCTGATCCGAGCTTGGAGGCC 101

RESULT 12
AQ535244 103 bp DNA GSS 18-MAY-1999
LOCUS RPCI-11-31/H22.TV RPCI-11 Homo sapiens genomic clone
DEFINITION RPCI-11-31/H22, genomic survey sequence.
ACCESSION AQ535244
VERSION AQ535244.1 GI:4846934
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and
Venter, J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building

JOURNAL COMMENT Unpublished (1997)
Contact: Shaying Zhao, William Niernman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetlgr.org

Clones are derived from the human BAC library RPC1-11. For BAC library availability, please contact Pieter de Jong (pieter@ejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genet cs (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html.
Seq primer: T7
Class: BAC ends.

FEATURES

source

1..103
/organism="Homo sapiens"
/db_xref="GDB:7621533"
/db_xref="taxon:9606"
/clone="RPC1-11-31H2"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC library"

BASE COUNT 31 a 27 c 27 g 18 t

ORIGIN

Query Match 0.1%; Score 37; DB 104; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7322 AGAGACGGGTTTCACCATGTTGGCCAGCGTGTCTC 7358
|||||
Db 87 AGAGACGGGTTTCACCATGTTGGCCAGCGTGTCTC 51

RESULT 13 87 bp mRNA EST 30-JUL-1996
D20989/c LOCUS HUMGS01971 Human promyelocyte Homo sapiens cDNA clone mp0383 3',
DEFINITION mRNA sequence.
ACCESSION D20989
VERSION D20989.1 GI:504809
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 87)
Okubo,K., Fukushima,A., Yoshii,J., Miyama,T., Kojima,Y.,

Yoshinari,H., Arimoto,J. and Matsubara,K.

Gene expression of human promyelocytic cell line HL60 before and

after induction of differentiation. A new application of 3'directed

cDNA sequencing

Unpublished (1993)

Contact: Okubo,K., Fukushima,A., Yoshii,J., Miyama,T., Kojima,Y.,

Yoshinari,H., Arimoto,J. and Matsubara,K.

Institute for Molecular and Cellular Biology

Osaka University

3-1 Yamada-oka,Suita,Osaka 565,Japan.

Location/Qualifiers

FEATURES

source

1..87
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="mp0383"
/clone_1lb="Human promyelocyte"
/note="Female; adult; cell_line = HL60; cell_type =
promyelocyte."

BASE COUNT 13 a 23 c 17 g 34 t

ORIGIN

Query Match 0.1%; Score 36; DB 20; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.24;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22232 GTGGCTACACCTGTATCCAGCACTTGGGAGGC 22267

Db 51 GTGGCTACACCTGTATCCAGCACTTGGGAGGC 16

RESULT 14

AA464896/c LOCUS AA464896

DEFINITION aa39e05.r1 Striatogene fetal retina 937202 Homo sapiens cDNA clone

IMAGE:838880 5' similar to contains Alu repetitive element; mRNA

sequence.

ACCESSION AA464896

VERSION AA464896

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 97)
Hiller,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,

Kritman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,

Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,

Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.

WashU-NCI human EST Project

Unpublished (1997)

On May 5, 1995 this sequence version replaced gi:798258.

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert length: 758 Std Error: 0.00

Seq primer: -28m3 rev1 ET from Amersham

High quality sequence stop: 97.

FEATURES

source

1..97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:838880"
/clone_1lb="Striatogene fetal retina 937202"
/sex="mixed"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled
retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
Vector: -5' adaptor sequence: 5' GAATTCGGACGCG 3' -3'
adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3' "

BASE COUNT 17 a 25 c 26 g 29 t

ORIGIN

Query Match 0.1%; Score 36; DB 34; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22374 GTGCCGTATCCAGCACTACTAGAGAGCTGAGCA 22409

Db 38 GTGCCGTATCCAGCACTACTAGAGAGCTGAGCA 3

RESULT 15

B36140/c LOCUS B36140 65 bp DNA GSS 17-OCT-1997

Mon Apr 3 08:24:05 2000

us-08-852-495c-1_copy_115000_145000.rst

Page 8

DEFINITION HS-1038-A1-D06-MR.ab1 CIT Human Genomic Sperm Library C Homo
sapiens genomic clone Plate-CT 820 Col=11 Row=G, genomic survey
sequence.
ACCESSION B36140
VERSION B36140.1 GI:2535509
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 65)
AUTHORS Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
Tratcoff,R., Adjajian,C., Blanchard,A., West,A. and Hood,L.E.
TITLE Construction of a Characterized Clone Resource for Genomic
Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
Tagged Connectors
JOURNAL Unpublished (1997)
COMMENT Contact: Mahairas GG, Zackrone KD, Hood L
University of Washington
Seattle, WA 98195, USA
Tel: (206) 616-8744
Fax: (206) 685-7301
Email: kzackron@u.washington.edu
Sequence Tagged Connector
Plate: CT 820 row: G column: 11
Class: BAC ends
High quality sequence stop: 65.
FEATURES
source
1
65
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-CT 820 Col=11 Row=G"
/clone_lib="CIT Human Genomic Sperm Library C"
/sex="M"
/note="Organ: sperm; Vector: pBelbac11; BAC clones in
E-Coli DH10B"
BASE COUNT 25 a 18 c 10 g 12 t
ORIGIN
Query Match 0.1%; Score 35; DB 81; Length 65;
Best local similarity 100.0%; Pred. No. 0.59;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16455 AGACGGGGTTTCACCATGTTGTCAGGCTGCTC 16489
|||||
DB 35 AGACGGGGTTTCACCATGTTGTCAGGCTGCTC 1
Search completed: March 30, 2000, 19:57:22
Job time: 146496 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 23:41:00 ; Search time 13413 Seconds

(without alignments)
-6791.518 Million cell updates/sec

Title: US-08-852-495c-1_COPY_140000_170000

Perfect score: 30001

Sequence: 1 CCCGCCAATCCCATATGCAC.....TACATATTTATATATGTA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 821193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database :

GenDb1:*
1: gb_bal:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_p11:*
8: gb_p12:*
9: gb_p13:*
10: gb_p14:*
11: gb_p15:*
12: gb_p16:*
13: gb_p17:*
14: gb_p18:*
15: gb_p19:*
16: gb_p20:*
17: gb_p21:*
18: gb_p22:*
19: gb_p23:*
20: gb_p24:*
21: gb_p25:*
22: gb_p26:*
23: gb_p27:*
24: gb_p28:*
25: gb_p29:*
26: gb_p30:*
27: gb_p31:*
28: gb_p32:*
29: gb_p33:*
30: gb_p34:*
31: gb_p35:*
32: gb_p36:*
33: gb_p37:*
34: gb_p38:*
35: gb_p39:*
36: gb_p40:*
37: gb_p41:*
38: gb_p42:*
39: gb_p43:*
40: gb_p44:*
41: gb_p45:*
42: gb_p46:*
43: gb_p47:*

44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_p13:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	0.1	80	9 HUMBRKFAA	M36131 Human alpha
2	42	0.1	80	9 HUMBRKFAA	M36132 Human alpha
3	40	0.1	40	5 A68622	A68622 Sequence 2
4	39	0.1	90	5 HUMLDLRF1	K03555 Human low d
5	38	0.1	60	5 AR051487	AR051487 Sequence
6	38	0.1	76	5 AR051499	AR051499 Sequence
7	38	0.1	80	9 HUMBRKFAA	M36133 Human alpha
8	38	0.1	90	9 HUMLDLRF1	M15365 Human low d
9	38	0.1	97	9 HUMLDLRF2	M14180 Human low d
10	37	0.1	97	9 HUMLDLRF3	M14179 Human fam1
11	37	0.1	101	10 S79560	S79560 HRX (Inton
12	37	0.1	101	10 S79561	S79561 dHRX (partl
13	36	0.1	40	5 A68621	A68621 Sequence 1
14	36	0.1	51	9 HUMDB603M3	D17279 Human HepG2
15	36	0.1	103	9 HUMALCE221	M87896 Human carcl
16	36	0.1	105	13 G32655	G32655 A009130 Hum
17	35	0.1	35	5 A25212	A25212 Inter-Alu S
18	35	0.1	35	5 E09140	E09140 Synthetic D
19	35	0.1	70	9 HSLAS32	X91547 H. sapiens D
20	35	0.1	91	13 HUMUT8164A	L30244 Human STS U
21	34	0.1	63	13 HUMUT5302A	L30829 Human STS U
22	34	0.1	80	9 HUMBRKFAA	M36132 Human alpha
23	34	0.1	80	9 HUMBRKFAA	M36133 Human alpha
24	33	0.1	80	9 HUMBRKFAA	M36135 Human alpha
25	33	0.1	100	9 HUMGALNSA	D45223 Human GALNS
26	32	0.1	51	10 S62605	S62605 C1-Inhibito
27	31	0.1	60	13 HUMUT243A	L29936 Human STS U
28	31	0.1	67	9 HUMALINCD	L36843 Homo sapien
29	31	0.1	77	13 HUMUT6154A	L30742 Human STS U
30	31	0.1	77	13 HUMUT6154A	L30742 Human STS U
31	31	0.1	90	9 HUMLDLRF1	K03556 Human low d
32	31	0.1	97	9 HUMLDLRF1	M14178 Human low d
33	31	0.1	104	9 HUMALCE272	M87899 Human carcl
34	30	0.1	50	10 S62604	S62604 C1-Inhibito
35	30	0.1	63	13 HUMUT5302A	L30829 Human STS U
36	30	0.1	91	13 HUMUT8164A	L30244 Human STS U
37	30	0.1	95	4 GGER10H7	X78616 G. gallus ge
38	29	0.1	99	13 HUMUT7692A	L30306 Human STS U
39	29	0.1	30	5 AR051440	AR051440 Sequence
40	29	0.1	84	5 AR051521	AR051521 Sequence
41	28	0.1	28	5 A49272	A49272 Sequence 2
42	28	0.1	50	5 I23510	I23510 Sequence 5
43	28	0.1	50	5 I28359	I28359 Sequence 5
44	28	0.1	50	5 I28514	I28514 Sequence 5
45	28	0.1	50	5 I41125	I41125 Sequence 5

ALIGNMENTS

RESULT 1
HUMBRKFAA/c
LOCUS HUMBRKFAA 80 bp DNA
DEFINITION Human alpha-galactosidase breakpoint region.
ACCESSION M36131
VERSION M36131.1 GI:179541

27-APR-1993

KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.

SOURCE Human DNA.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS 1 (bases 1 to 80)

TITLE Kornreich, R., Bishop, D.F. and Desnick, R.J.

JOURNAL Alpha-galactosidase A gene rearrangements causing Fabry disease: identification of short direct repeats at breakpoints in an Alu-rich gene

MEDLINE J. Biol. Chem. 265, 9319-9326 (1990)

FEATURES Location/Qualifiers

source 1..80

BASE COUNT 19 a 28 c 20 g 13 t

ORIGIN

Query Match 0.1%; Score 44; DB 9; Length 80;

Best Local Similarity 100.0%; Pred. No. 2.8e-10;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13518 GTGGGCTCATGCTGTATCCAGCAGCTTTGGAGGCCAAG 13561

Db 60 GTGGGCTCATGCTGTATCCAGCAGCTTTGGAGGCCAAG 17

RESULT 2

LOCUS HMBKRFAB 80 bp DNA PRI 27-APR-1993

DEFINITION Human alpha-galactosidase breakpoint family F.

ACCESSION M36132

VERSION M36132.1 GI:179542

KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.

SOURCE Human DNA.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS 1 (bases 1 to 80)

TITLE Kornreich, R., Bishop, D.F. and Desnick, R.J.

JOURNAL Alpha-galactosidase A gene rearrangements causing Fabry disease: identification of short direct repeats at breakpoints in an Alu-rich gene

MEDLINE J. Biol. Chem. 265, 9319-9326 (1990)

FEATURES Location/Qualifiers

source 1..80

BASE COUNT 17 a 29 c 16 g 18 t

ORIGIN

Query Match 0.1%; Score 42; DB 9; Length 80;

Best Local Similarity 100.0%; Pred. No. 2.7e-09;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13520 GTGGGCTCATGCTGTATCCAGCAGCTTTGGAGGCCAAG 13561

Db 58 GTGGGCTCATGCTGTATCCAGCAGCTTTGGAGGCCAAG 17

RESULT 3

LOCUS A68622/c 40 bp DNA PAT 06-MAY-1999

DEFINITION Sequence 2 from Patent WO9801573.

ACCESSION A68622

VERSION A68622.1 GI:4759649

KEYWORDS

SOURCE unidentified.

ORGANISM unidentified

REFERENCE unclassified.

AUTHORS 1 (bases 1 to 40)

TITLE Resnick, M.A., Laktionov, V.L., Kornreich, N.Y. and Perkins, E.L.

JOURNAL TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING

US HEALTH (US) 15-JAN-1998;

US HEALTH (US) 15-JAN-1998;

FEATURES Location/Qualifiers

source 1..40

BASE COUNT 9 a 8 c 19 g 4 t

ORIGIN

Query Match 0.1%; Score 40; DB 5; Length 40;

Best Local Similarity 100.0%; Pred. No. 2.7e-08;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28376 CCGCTCCCGGTTCAAGCATTCCTGCTCAGCTCC 28415

Db 40 CCGCTCCCGGTTCAAGCATTCCTGCTCAGCTCC 1

RESULT 4

LOCUS HMBKRFAB/c 90 bp DNA PRI 11-JAN-1995

DEFINITION Human low density lipoprotein receptor intron A Alu repeat.

ACCESSION K03555

VERSION K03555.1 GI:187104

KEYWORDS Alu repeat; low density lipoprotein receptor-1; repeat region.

SOURCE Human DNA.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS 1 (bases 1 to 90)

TITLE Lehman, M.A., Goldstein, J.L., Russell, D.W. and Brown, M.S.

JOURNAL duplication of seven exons in LDL receptor gene caused by Alu-Alu recombination in a subject with familial hypercholesterolemia

MEDLINE Cell 48 (5), 827-835 (1987)

COMMENT Clean copy of sequence kindly provided by M. Lehman (22-APR-1987).

FEATURES Location/Qualifiers

source 1..90

BASE COUNT 17 a 30 c 24 g 19 t

ORIGIN

Query Match 0.1%; Score 39; DB 9; Length 90;

Best Local Similarity 100.0%; Pred. No. 8.2e-08;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13519 TGGTGCTCATGCTGTATCCAGCAGCTTTGGAGGCC 13557

Db 80 TGGTGCTCATGCTGTATCCAGCAGCTTTGGAGGCC 42

RESULT 5

LOCUS ARO51487 60 bp DNA PAT 29-SEP-1999

DEFINITION Sequence 57 from patent US 5830670.

ACCESSION ARO51487

VERSION ARO51487.1 GI:5974851

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE unclassified.

1 (bases 1 to 60)


```
AUTHORS de la Monte,S. and Wands,J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's
JOURNAL Patent: US 5830670-A 57 03-NOV-1998;
FEATURES Location/Qualifiers
SOURCE 1..60
BASE COUNT 12 a 14 c 15 g 19 t
ORIGIN

Query Match 0.1%; Score 38; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9496 GCTAATTTTGTATTAGTAGAGACAGGGTTTCACC 9533
Db 11 GCTAATTTTGTATTAGTAGAGACAGGGTTTCACC 48

RESULT 6
LOCUS AR051499 76 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 69 from patent US 5830670.
ACCESSION AR051499
VERSION AR051499.1 GI:5974863
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 76)
AUTHORS de la Monte,S. and Wands,J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's
JOURNAL Patent: US 5830670-A 69 03-NOV-1998;
FEATURES Location/Qualifiers
SOURCE 1..76
BASE COUNT 15 a 19 c 18 g 24 t
ORIGIN

Query Match 0.1%; Score 38; DB 5; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9496 GCTAATTTTGTATTAGTAGAGACAGGGTTTCACC 9533
Db 11 GCTAATTTTGTATTAGTAGAGACAGGGTTTCACC 48

RESULT 7
LOCUS HUMBRKFA2/c 80 bp DNA PRI 27-APR-1993
DEFINITION Human alpha-galactosidase breakpoint region.
ACCESSION M36133
VERSION M36133.1 GI:179543
KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
Kornreich,R., Bishop,D.F. and Desnick,R.J.
TITLE Alpha-galactosidase A gene rearrangements causing Fabry disease:
AUTHORS Identification of short direct repeats at breakpoints in an
JOURNAL Alu-rich gene
MEDLINE J. Biol. Chem. 265, 9319-9326 (1990)
FEATURES Location/Qualifiers
SOURCE 1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 18 a 28 c 17 g 17 t
ORIGIN

Query Match 0.1%; Score 38; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13520 GGTGGCTCATGCTGTATATCCACGACATTGGAGGCC 13557
Db 58 GGTGGCTCATGCTGTATATCCACGACATTGGAGGCC 21

RESULT 8
LOCUS HUMDLR2/c 90 bp DNA PRI 11-JAN-1995
DEFINITION Human low density lipoprotein receptor mutant gene recombination
site.
ACCESSION M15365
VERSION M15365.1 GI:187107
KEYWORDS Alu repeat; LDL receptor; cell surface protein; crossover;
recombination.
SOURCE Human (FH 295) fibroblast DNA, clone p295.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 90)
AUTHORS Lehman,M.A., Goldstein,J.L., Russell,D.W. and Brown,M.S.
TITLE Duplication of seven exons in LDL receptor gene caused by Alu-Alu
JOURNAL recombination in a subject with familial hypercholesterolemia
MEDLINE Cell 48 (5), 827-835 (1987)
COMMENT 8731094
Clean copy of sequence kindly provided by M. Lehman (22-APR-1987).
Individual FH 295 carries two mutant LDL receptor alleles. The
allele below was inherited from the father and includes a
duplication of exons 2 through 8 that was presumably created by
unequal chromosomal crossover involving Alu repeats in introns 1
and 8. The exact site of recombination cannot be determined
because the sequences of the Alu repeats in introns 1 and 8 of
normal alleles are identical over a span of 26 nucleotides at the
recombination site.
FEATURES Location/Qualifiers
SOURCE 1..90
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19p13.3"
1..43
/gene="LDLR"
<1..43
/gene="LDLR"
/note="LDLR Intron 8; G00-119-362"
43..68
/organism="Homo sapiens"
68..>90
/note="LDLR duplicated Intron 1 (no splice consensus at
58); putative; does not fit consensus"
BASE COUNT 18 a 33 c 19 g 20 t
ORIGIN 1 bp upstream of NheI site; chromosome 19p13.2-p13.1.

Query Match 0.1%; Score 38; DB 9; Length 90;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13519 TGGTGGCTCATGCTGTATATCCACGACATTGGAGGCC 13556
Db 80 TGGTGGCTCATGCTGTATATCCACGACATTGGAGGCC 43

RESULT 9
LOCUS HUMDLR2/c 97 bp DNA PRI 07-JAN-1995
DEFINITION Human low density lipoprotein receptor gene (LDLR), Intron 5
```

```
(partial).
ACCESSION      M14180
VERSION        M14180.1  GI:187098
KEYWORDS       low density lipoprotein receptor-1.
SEGMENT        2 of 2
SOURCE         Human white blood cell DNA.
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
                Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 97)
AUTHORS        Hobbs, H.H., Brown, M.S., Goldstein, J.L. and Russell, D.W.
TITLE          Deletion of exon encoding cysteine-rich repeat of low density
                lipoprotein receptor alters its binding specificity in a subject
                with familial hypercholesterolemia
JOURNAL        J. Biol. Chem. 261 (28), 13114-13120 (1986)
MEDLINE        87008518
COMMENT        Analysis of the LDL-receptor gene of a patient with familial
                hypercholesterolemia (FH) revealed the deletion of exon 5 resulting
                from a homologous recombination between repetitive Alu sequences of
                intron 4 and intron 5.
FEATURES       source
                1..97
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /map="19p13.3"
                join(M14178..11..97,1..97)
                /gene="LDLR"
                <1..>97
                /gene="LDLR"
                /note="LDLR intron E; G00-119-362"
                42..72
                /gene="LDLR"
                /note="deletion target sequence"
                34 c 25 g 19 t
                /misc_feature
                19 a 34 c 25 g 19 t
                /ORIGIN
                Chromosome 19p13.2-p13.1.

Query Match
Best Local Similarity 100.0%; Score 38; DB 9; Length 97;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13529 TGCGTGTATCCAGCAGCTTGGAGGCCAGCGAGC 13566
Db 79 TGCGTGTATCCAGCAGCTTGGAGGCCAGCGAGC 42

RESULT 10
LOCUS          HUMLDLRD/c 97 bp DNA PRI 27-APR-1993
DEFINITION     Human familial hypercholesterolemia 626-a gene with a deletion of
                exon 5.
ACCESSION      M14179
VERSION        M14179.1  GI:187101
KEYWORDS       low density lipoprotein receptor-1; very low density lipoprotein.
SOURCE         Human fibroblast DNA, from patient FH-626.
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
                Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 97)
AUTHORS        Hobbs, H.H., Brown, M.S., Goldstein, J.L. and Russell, D.W.
TITLE          Deletion of exon encoding cysteine-rich repeat of low density
                lipoprotein receptor alters its binding specificity in a subject
                with familial hypercholesterolemia
JOURNAL        J. Biol. Chem. 261, 13114-13120 (1986)
MEDLINE        87008518
COMMENT        Analysis of the LDL-receptor gene of a patient with familial
                hypercholesterolemia (FH) revealed the deletion of exon 5 resulting
                from a homologous recombination between repetitive Alu sequences of
                intron 4 and intron 5.
FEATURES       source
                1..97
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /db_xref="taxon:9606"

misc_feature 42..72
                /note="deletion target sequence"
                34 c 23 g 18 t
                /ORIGIN
                Chromosome 19p13.2-p13.1.

Query Match
Best Local Similarity 100.0%; Score 37; DB 10; Length 101;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9514 AGTAGAGCAGGAGTTCCACCATGTGGCGAGGCTGT 9550
Db 50 AGTAGAGCAGGAGTTCCACCATGTGGCGAGGCTGT 14

RESULT 12
LOCUS          S79561/c 101 bp DNA PRI 27-JAN-1996
DEFINITION     dHXX (partial genomic duplication startpoint) [human, acute myeloid
                leukemia with trisomy 11 syndrome patient J, genomic Mutant, 101
                nt].
ACCESSION      S79561
VERSION        S79561.1  GI:1168042
KEYWORDS       human acute myeloid leukemia with trisomy 11 syndrome patient J.
SOURCE         Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
                Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 101)
AUTHORS        Bernard, O.A., Romana, S.P., Schichman, S.A., Mauchauffe, M.,
                Jonveaux, P. and Berger, R.
TITLE          Partial duplication of HRX in acute leukemia with trisomy 11
                Map location: 11q23.
FEATURES       source
                1..101
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                1..101
                /partial
                /gene="HRX"
                /ORIGIN
                27 a 21 c 28 g 25 t

BASE COUNT    27 a 21 c 28 g 25 t
ORIGIN
Chromosome 11q23.

Query Match
Best Local Similarity 100.0%; Score 37; DB 10; Length 101;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9514 AGTAGAGCAGGAGTTCCACCATGTGGCGAGGCTGT 9550
Db 50 AGTAGAGCAGGAGTTCCACCATGTGGCGAGGCTGT 14

RESULT 12
LOCUS          S79561/c 101 bp DNA PRI 27-JAN-1996
DEFINITION     dHXX (partial genomic duplication startpoint) [human, acute myeloid
                leukemia with trisomy 11 syndrome patient J, genomic Mutant, 101
                nt].
ACCESSION      S79561
VERSION        S79561.1  GI:1168042
KEYWORDS       human acute myeloid leukemia with trisomy 11 syndrome patient J.
SOURCE         Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
                Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 101)
AUTHORS        Bernard, O.A., Romana, S.P., Schichman, S.A., Mauchauffe, M.,
                Jonveaux, P. and Berger, R.
TITLE          Partial duplication of HRX in acute leukemia with trisomy 11
                Map location: 11q23.
FEATURES       source
                1..101
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                1..101
                /partial
                /gene="HRX"
                /ORIGIN
                27 a 21 c 28 g 25 t

BASE COUNT    27 a 21 c 28 g 25 t
ORIGIN
Chromosome 11q23.
```

JOURNAL
MEDLINE
REMARK
Leukemia 9 (9), 1487-1490 (1995)
Genbank staff at the National Library of Medicine created this entry [NCBI gblsbs 170936] from the original journal article. This sequence comes from Fig. 3.
Map location: 11q23.
Authors indicate breakpoint region at 48-62.
Location/Qualifiers
1. .101
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. .101
/partial
/gene="dhrx"
22 a 22 c 23 g 34 t

BASE COUNT
ORIGIN

Query Match 0.1%; Score 37; DB 10; Length 101;
Best Local Similarity 100.0%; Pred. No. 8e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9514 AGTAGAGACGGCTTCACCATGTGGCCAGCGTGT 9550
DB 50 AGTAGAGACAGGGTTTCACCATGTGGCCAGCGTGT 14

RESULT 13
LOCUS A68621 40 bp DNA
DEFINITION Sequence 1 from Patent WO9801573.
ACCESSION A68621
VERSION A68621.1 GI:4759648
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Resnick, M.A., Laktionov, V.L., Kouprina, N.Y. and Perkins, E.L.
TITLE TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
JOURNAL Patent: WO 9801573-A 15-JAN-1998;
US HEALTH (US)

FEATURES
source location/Qualifiers
1. .40
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 7 a 12 c 13 g 8 t
ORIGIN

Query Match 0.1%; Score 36; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24026 GTGGCTACCCCTGTATCCACAGACTTTGGAGGC 24061
DB 2 GTGGCTACCCCTGTATCCACAGACTTTGGAGGC 37

RESULT 14
LOCUS HUMD6B03M3 51 bp mRNA
DEFINITION Human Hepg2 3' region MboI cDNA, clone hmd6b03m3.
ACCESSION D17279
VERSION D17279.1 GI:598922
KEYWORDS gene signature.
SOURCE Homo sapiens male cell_line:Hepg2 cDNA to mRNA, clone_11b:Kiseru.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 51)
AUTHORS Matoba, R.
TITLE Direct Submission

JOURNAL
MEDLINE
REMARK
Submitted (21-Jul-1993) to the DDBJ/EMBL/Genbank databases. Ryo Matoba, Osaka University, Institute for Molecular and Cellular Bio; 1-3 Yamada-oka, Suita, Osaka 565, Japan
E-mail: matoba@inherl.lnmb.osaka-u.ac.jp;
Tel: 81-6-877-5111 (ex. 3314), Fax: 81-6-877-1922)
2 (bases 1 to 51)
Matoba, R., Okubo, K., Hori, N., Fukushima, A. and Matsubara, K.
The addition of 5'-coding information to a 3'-directed cDNA library improves analysis of gene expression
Gene 146 (2), 199-207 (1994)
94357437
Submitted (21-Jul-1993) to DDBJ by:
Ryo Matoba
Molecular Microbiology and Genetics Lab.
Research Institute of Innovative Technology for the Earth 9-2
Kizugawadal Kizu-cho,
Soraku-gun, Kyoto
Japan, 619-02
Phone: 07747-5-2308
Fax: 07747-5-2321.

FEATURES
source location/Qualifiers
1. .51
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HepG2"
/clone_11b="Kiseru"
/sex="Male"

BASE COUNT 13 a 19 c 12 g 7 t
ORIGIN

Query Match 0.1%; Score 36; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19138 CCACCTGACTCCAGCCTGGCGCAGAGGAGACTC 19173
DB 8 CCACCTGACTCCAGCCTGGCGCAGAGGAGACTC 43

RESULT 15
LOCUS HUMACE221/c 103 bp ss-RNA
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Simmet, D., Richer, C., Deragon, J.-M. and Labuda, D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) in press
FEATURES
source location/Qualifiers
1. .103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HepG2"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"

BASE COUNT 25 a 27 c 33 g 18 t
ORIGIN

Query Match 0.1%; Score 36; DB 9; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28372 ACCTCCGCTCCCGGGTTCAAGCATTCCTGCT 28407

Mon Apr 3 08:24:06 2000

us-08-852-495c-1_copy_140000_170000.rge

Page 6

Db 69 ACCCTCGGCTCCCGGGTTCAAGGATCTCTCGCT 34

Search completed: March 31, 2000, 06:36:34
Job time: 184343 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2000, 00:01:30 ; Search time 600.33 Seconds
(without alignments)
12503.144 Million cell updates/sec

Title: US-08-852-495c-1_COPY_140000_170000
Perfect score: 30001
Sequence: 1 CCGTCCAAATCCATATGCAC.....TACATATTTATATATGTA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 311585 seqs, 125096042 residues

Word size : 0

Total number of hits satisfying chosen parameters: 433264

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	40	0.1	40	1	Alu PCR primer 2.
C 2	36	0.1	40	1	Alu PCR primer 1.
C 3	35	0.1	35	1	Inter-Alu specific
C 4	29	0.1	30	1	Neutral thread prot
C 5	29	0.1	30	1	Neutral thread prot
C 6	29	0.1	55	1	CDNA synthesis pri
C 7	29	0.1	58	1	CDNA synthesis pri
C 8	28	0.1	28	1	Primer Alu B corre
C 9	28	0.1	32	1	Enhancer element e
C 10	28	0.1	32	1	Enhancer element e
C 11	28	0.1	40	1	Sequence of primer
C 12	28	0.1	48	1	Trypsin inhibitor
C 13	28	0.1	49	1	Human secreted pro
C 14	28	0.1	49	1	3' portion of cDNA
C 15	28	0.1	53	1	DNA detector seque
C 16	28	0.1	60	1	PCR primer Alu. Nu
C 17	28	0.1	86	1	Self-complementary
C 18	28	0.1	98	1	Porcine reproductive
C 19	28	0.1	99	1	Split oligonucleo
C 20	28	0.1	101	1	polyA-LTR oligonuc
C 21	27	0.1	29	1	PR-1 promoter prim
C 22	27	0.1	36	1	PR-1 promoter prim
C 23	27	0.1	41	1	Aspergillus oryzae
C 24	27	0.1	43	1	Oligonucleotide 43
C 25	27	0.1	43	1	Oligonucleotide 43
C 26	27	0.1	47	1	Template switching
C 27	27	0.1	48	1	Intronic human MSH
C 28	27	0.1	50	1	Primer, RI-Not-T30
C 29	27	0.1	50	1	Primer, MBTC from w
C 30	27	0.1	50	1	PCR suppression pr
C 31	27	0.1	50	1	Primer of the spec
C 32	27	0.1	50	1	EP-892047 Seq ID 1
C 33	27	0.1	50	1	CDNA synthesis pri

C 34	27	0.1	51	1	T98497
C 35	27	0.1	52	1	T33764
C 36	27	0.1	52	1	T98502
C 37	27	0.1	52	1	T98005
C 38	27	0.1	52	1	V24462
C 39	27	0.1	52	1	V24019
C 40	27	0.1	52	1	V35496
C 41	27	0.1	52	1	X19061
C 42	27	0.1	53	1	V83890
C 43	27	0.1	54	1	T17031
C 44	27	0.1	54	1	T72939
C 45	27	0.1	55	1	Q36440

ALIGNMENTS

RESULT 1
V19045/c
ID V19045 standard; DNA: 40 BP.
AC V19045;
DT 28-JUL-1998 (first entry)
DE Alu PCR primer 2; Affiliation: Alu repeat sequence; vector;
KW PCR; primer; amplification; Alu repeat sequence; vector;
KW circular yeast artificial chromosome; YAC; ss.
OS Synthetic.
OS Saccharomyces sp.
PN MO9801573-Al.
PD 15-JAN-1998.
PR 09-JUL-1996; U11478.
DR WPI: 98-110234/10.
PT Preparation of yeast artificial chromosomes - by in vivo
recombination using vector comprising yeast centromere, marker,
PS yeast telomere and nucleic acid for recombination
PT Example 1; Page 45; 117pp; English.
CC This is the nucleotide sequence for the PCR primer used in the
CC amplification of the Alu repeat sequence, which is used to
CC demonstrate the processes described in the invention. It involves
CC the creation and use of circular yeast artificial chromosome (YAC)
CC to selectively clone specific nucleic acids from a background of
CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
CC They can be used to rapidly isolate human DNA where only a part of the
CC sequence of DNA is known. Using the methods large fragments of DNA can
CC be easily cloned and analysed.
SQ Sequence 40 BP; 9 A; 8 C; 19 G; 4 T;

Query Match 0.1%; Score 40; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28376 CCGCTCCCGGCTCAAGCATTCCTCCAGCCTCC 28415
DB 40 CCGCTCCCGGCTCAAGCATTCCTCCAGCCTCC 1

RESULT 2
V19044
ID V19044 standard; DNA: 40 BP.
AC V19044;
DT 28-JUL-1998 (first entry)
DE Alu PCR primer 1.
KW PCR; primer; amplification; Alu repeat sequence; vector;
KW circular yeast artificial chromosome; YAC; ss.
OS Synthetic.
OS Saccharomyces sp.
PN MO9801573-Al.
PD 15-JAN-1998.
PR 09-JUL-1996; U11478.
DR WPI: 98-110234/10.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI KOPRISHA NY LARIONOV VL, FERRINS EL, KESNICK MA;
DR WPI: 36-110234/10.
PT Preparation of yeast artificial chromosomes - by in vitro
PT recombination using vector comprising yeast centromere, marker,
PT yeast telomere and nucleic acid for recombination
PS Example 1; Page 45; 11/9p; English.
CC This is the nucleotide sequence for the PCR primer used in the
CC amplification of the Alu repeat sequence, which is used to
CC demonstrate the processes described in the invention. It involves
CC the creation and use of circular yeast artificial chromosome (YAC)
CC to selectively clone specific nucleic acids from a background of
CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
CC They can be used to rapidly isolate human DNA where only a part of the
CC sequence of DNA is known. Using the methods large fragments of DNA can
CC be easily cloned and analysed.
SQ Sequence 40 BP; 7 A; 12 C; 13 G; 8 T;

Query Match 0.1%; Score 36; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.36; Mismatches 0; Gaps 0;
Matches 36; Conservative 0; Indels 0;

OY 24026 GTGGCTACGCGCTGTATCCAGCACTTGGAGGC 24061

DB 2 GTGGCTACGCGCTGTATCCAGCACTTGGAGGC 37

RESULT 3
OY 24026 standard; DNA: 35 BP.
AC Q27391.
DE 27-JAN-1993 (first entry)
KW Inter-Alu specific primer PDJ33.
KW Polymerase chain reaction; PCR; repetitive element; ss.
OS Synthetic.
PN WO9213101-A.
PD 06-AUG-1992.
PF 24-JAN-1992; NL0018.
PR 25-JAN-1991; NL-000132.
PA (INGE-) INGENY BV.
PI Uiterlinden AG, Vijg J;
DR WPI: 92-284683/34.
PT Detection of genetic variation by 2-D electrophoresis of
PT fragments - and hybridisation with labelled probes, carried out
PT on fragments consisting of inter-repeat sequences generated by
PT PCR.
PS Claim 6; Page 6; 31pp; English.
CC Primer PDJ33 is one of several primers which are preferred for use
CC in amplifying inter-Alu regions of DNA. The amplified fragments are
CC then subjected to 2-D electrophoresis on the basis of length and
CC differences in base sequence. The resulting separation pattern is
CC transferred to a filter for screening with a probe. The method can
CC be used to detect genetic variation.
CC See Q27389-Q27404 and Q23111-Q23144.
SQ Sequence 35 BP; 8 A; 10 C; 11 G; 6 T;

Query Match 0.1%; Score 35; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.69; Mismatches 0; Gaps 0;
Matches 35; Conservative 0; Indels 0;

OY 24027 TGGCTACGCGCTGTATCCAGCACTTGGAGGC 24061

DB 35 TGGCTACGCGCTGTATCCAGCACTTGGAGGC 1

RESULT 4
OY 24027 standard; CDNA: 30 BP.
AC Q27890.
DE 06-JUL-1995 (first entry)
PI Neutral thread protein AD10-7 CDNA 5' antisense oligonucleotide.

KW Neutral thread protein AD10-7; Alzheimer's; neuroectodermal tumours;
KW malignant astrocytomas; glioblastomas; 5' antisense therapy; ss.
OS Synthetic.
PN WO9423756-A.
PD 27-OCT-1994.
PF 20-APR-1994; US-04321.
PR 20-APR-1993; US-050559.
PA (GHEO) GEN HOSPITAL CORP.
PI De LA MONTE SM, Wands JR;
DR WPI: 94-341497/42.
PT Detection of neural thread proteins - to detect sporadic and
PT familial Alzheimer's disease, neuroectodermal tumours, malignant
PT astrocytomas and glioblastomas (Eng).
PS Disclosure; Page 48; 18pp; English.
CC Q27888-Q27890 are AD10-7 neutral thread protein (NTP) antisense
CC oligonucleotides, that can be used to down regulate or inhibit
CC the expression of the NTP gene. These oligonucleotides could be
CC used in the treatment of the following conditions Alzheimer's
CC disease, neuroectodermal tumours, malignant astrocytomas and
CC glioblastomas.
SQ Sequence 30 BP; 8 A; 4 C; 14 G; 4 T;

Query Match 0.1%; Score 29; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Gaps 0;
Matches 29; Conservative 0; Indels 0;

OY 28388 TTCAAGGATTCCTCGCTCAGCCTCC 28416

DB 30 TTCAAGGATTCCTCGCTCAGCCTCC 2

RESULT 5
OY 28388 standard; DNA: 30 BP.
AC T27744.
DE 14-NOV-1996 (first entry)
KW Neutral thread protein antisense sequence.
KW Neutral thread protein; NTP; diagnosis; detection;
KW Alzheimer's disease; neuroectodermal tumour; malignant astrocytoma;
KW monoclonal antibody; binding fragment; ds.
OS Synthetic.
PN WO9615272-A1.
PD 23-MAY-1996.
PF 14-NOV-1995; U17111.
PR 14-NOV-1994; US-340426.
PA (GHEO) GEN HOSPITAL CORP.
PI De LA MONTE S, Wands JR;
DR WPI: 96-259865/26.
PT Detection of neural thread protein in diagnosis of Alzheimer's
PT disease - also NTP DNA and protein sequences used in gene and
PT anti-sense therapy.
PS Disclosure; Page 48; 23pp; English.
CC A method for detecting the presence of neural thread protein (NTP)
CC having a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human
CC subject comprises (a) contacting a sample from a human subject that
CC is suspected of containing the NTP with at least one molecule
CC capable of binding to the protein, and (b) detecting any of the
CC from an antibody free of natural impurities, a monoclonal antibody
CC or a binding fragment of either of these. The method may be used for
CC diagnosing the presence of Alzheimer's disease, neuroectodermal
CC tumours and a malignant astrocytoma in a human. Expression of NTP
CC nucleic acid may be inhibited using antisense oligonucleotides
CC (See T27739-44).
SQ Sequence 30 BP; 8 A; 4 C; 14 G; 4 T;

Query Match 0.1%; Score 29; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 27; Mismatches 0; Gaps 0;
Matches 29; Conservative 0; Indels 0;

OY 28388 TTCAAGGATTCCTCGCTCAGCCTCC 28416

DB 30 TTCACGAGTCTCTGCGCCAGCCTCC 2
|||||
RESULT 6
ID T98499 standard; DNA; 55 BP.
AC T98499;
DT 11-MAR-1998 (first entry)
DE CDNA synthesis primer CDS3.
KW Template switching oligonucleotide; RNA/DNA hybrid; DNA preparation;
KW CDNA synthesis primer; CDNA cloning; CDNA library construction;
KW PCR primer; amplification; ss.
OS Synthetic.
PN WO9724455-A2.
PD 10-JUL-1997.
PF 03-JAN-1997; U00368.
PR 03-JAN-1996; US-582562.
PA (CLON-) CLONTECH LAB INC.
PI Chenchik A, Diatchenko L, Siebert P, Zhu Y;
DR WPI: 97-363690/33.
PT Preparation of cDNA from RNA molecules - by annealing CDNA synthesis
PT primer to RNA, synthesizing DNA and contacting with novel template
PT switching oligo:nucleotide
PS Claim 9; Page 29; 39pp; English.
CC T98497-T98520 represent primers used in the method of the invention. The
CC method of the invention is for preparing DNA complementary to the 5'-end
CC of an RNA molecule. The method comprises annealing a first DNA strand
CC primer to the RNA molecule and synthesizing a first DNA strand
CC complementary to at least a portion of the RNA molecule, and contacting
CC the RNA molecule with a template switching oligonucleotide (TSO) (such
CC as T98426) having a pre-selected nucleotide sequence at its 5'-end and at
CC least 1 riboguanine residue at its 3'-end, where the TSO binds the 5'-end
CC of the RNA molecule and serves as a template for the extension of the
CC 3'-end of the first DNA strand. The method can be used for the synthesis
CC and cloning of full length cDNA, or fragments, that correspond to the
CC complete sequence of the 5'-end of the mRNA molecule. It can be used to
CC construct cDNA libraries from nanogram quantities of total or poly A+
CC RNA. The TSO allows for negative selection against cDNA that are not
CC complementary to the 5'-end of the template RNA, while allowing full
CC length cDNA to be readily selected.
SQ Sequence 55 BP; 10 A; 5 C; 7 G; 33 T;

Query Match 0.1%; Score 29; DB 1; Length 55;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26680 AAAAAAAAAAAAAAAAAAGTA 26708
|||||
DB 51 AAAAAAAAAAAAAAAAAAGTA 23

RESULT 7
ID T98500 standard; DNA; 58 BP.
AC T98500;
DT 11-MAR-1998 (first entry)
DE CDNA synthesis primer Fr-T10NN.
KW Template switching oligonucleotide; RNA/DNA hybrid; DNA preparation;
KW CDNA synthesis primer; CDNA cloning; CDNA library construction;
KW PCR primer; amplification; ss.
OS Synthetic.
PN WO9724455-A2.
PD 10-JUL-1997.
PF 03-JAN-1997; U00368.
PR 03-JAN-1996; US-582562.
PA (CLON-) CLONTECH LAB INC.
PI Chenchik A, Diatchenko L, Siebert P, Zhu Y;
DR WPI: 97-363690/33.
PT Preparation of cDNA from RNA molecules - by annealing CDNA synthesis
PT primer to RNA, synthesizing DNA and contacting with novel template
PT switching oligo:nucleotide

PS Claim 9; Page 30; 39pp; English.
CC T98497-T98520 represent primers used in the method of the invention. The
CC method of the invention is for preparing DNA complementary to the 5'-end
CC of an RNA molecule. The method comprises annealing a first DNA strand
CC primer to the RNA molecule and synthesizing a first DNA strand
CC complementary to at least a portion of the RNA molecule, and contacting
CC the RNA molecule with a template switching oligonucleotide (TSO) (such
CC as T98426) having a pre-selected nucleotide sequence at its 5'-end and at
CC least 1 riboguanine residue at its 3'-end, where the TSO binds the 5'-end
CC of the RNA molecule and serves as a template for the extension of the
CC 3'-end of the first DNA strand. The method can be used for the synthesis
CC and cloning of full length cDNA, or fragments, that correspond to the
CC complete sequence of the 5'-end of the mRNA molecule. It can be used to
CC construct cDNA libraries from nanogram quantities of total or poly A+
CC RNA. The TSO allows for negative selection against cDNA that are not
CC complementary to the 5'-end of the template RNA, while allowing full
CC length cDNA to be readily selected.
SQ Sequence 58 BP; 8 A; 5 C; 9 G; 34 T;

Query Match 0.1%; Score 29; DB 1; Length 58;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26680 AAAAAAAAAAAAAAAAAAGTA 26708
|||||
DB 52 AAAAAAAAAAAAAAAAAAGTA 24

RESULT 8
ID T12509 standard; DNA; 28 BP.
AC T12509;
DT 09-SEP-1996 (first entry)
DE Primer A1u B corresp. to bases 294-267.
KW Primer: PCR; polymerase chain reaction; amplification; A1u repeat; ss;
KW quantitation; internal standard; plasmid; contamination; therapeutic.
OS Synthetic.
PN AT9401830-A.
PD 15-DEC-1995.
PF 26-SEP-1994; 001830.
PR 26-SEP-1994; AT-001830.
PA (IMMO) IMMUNO AG.
PI Dörner F, Falkner F, Haemmerle T, Himmelsbach M;
DR WPI: 96-069063/08.
PT Quantifying genomic DNA by amplification of repetitive sequences -
PT in presence of internal standard, then comparing amts of amplified
PT genomic DNA and standard, partic. for quality control of recombinant
PT viral proteins, vaccines, etc.
PS Claim 13; Page 14; 41pp; German.
CC Primers T12508-9 are used to amplify a 146 bp fragment of an A1u repeat
CC sequence in a novel method of quantitating genomic DNA in a sample. The
CC novel method features the addition of a known amount of an internal
CC standard nucleic acid which is different from the test nucleic acid in
CC at least one detectable feature. The template for these primers is the
CC plasmid pAlu-wt which comprises plasmid pCRII contg. nucleotides 148-294
CC of the A1u repeat sequence given in Nucleic Acid Res., 18 (1990) 6793.
CC The method is esp. useful for the detection of contaminating DNA in the
CC manufacture of therapeutic prods.
SQ Sequence 28 BP; 5 A; 9 C; 9 G; 5 T;

Query Match 0.1%; Score 28; DB 1; Length 28;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19152 CTTGGGCGACAGCGAGACTCTCTC 19179
|||||
DB 28 CTTGGGCGACAGCGAGACTCTCTC 1

RESULT 9

073570
ID 073570 standard; DNA: 32 BP.
AC 073570; (first entry)
DE Enhancer element er-3 conserved basepair sequence.
KW Enhancer element; carcinoma; tumor; cancer; SLPI gene;
KW secretory leukoprotease-inhibitor gene; cytokerin-8;
ss.
OS Homo sapiens.
FT Key Location/Qualifiers
FT misc.difference 29
FT /*tag- a
FT /*label- pyrimidine
PN W09421118-A.
PD 29-SEP-1994.
PE 24-MAR-1994; U03197.
PR 24-MAR-1993; US-035435.
PA (UABR-) UAB RES FOUND.
PI Garver RI, Sorscher EJ;
PI WPI: 94-316537/39.
DR DNA construct for treating human carcinoma - includes a
PT cancer-therapeutic gene under the control of a promoter and a gp.
PS Claim 1; Fig.6; 54pp; English.
CC This enhancer element is part of a DNA construct used for treating
CC human carcinoma which contains a cancer therapeutic protein under
CC the control of a promoter and 3 enhancer sequences in a specific 5'-3'
CC order. This enhancer element is derived from the flanking region of
CC the human epithelial cell cytokerin-8 gene.
SQ Sequence 32 BP; 7 A; 1 C; 8 G; 15 T;

Query Match 0.1%; Score 28; DB 1; Length 32;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28461 ATTTTGTATTATTAGATGAGTGGGT 28488
DB 1 ATTTTGTATTATTAGATGAGTGGGT 28

RESULT 10
ID 073570 standard; DNA: 32 BP.
AC 073570; (first entry)
DE Enhancer element er-3 conserved basepair sequence.
KW Enhancer element; carcinoma; tumor; cancer; SLPI gene;
KW secretory leukoprotease-inhibitor gene; cytokerin-8;
ss.
OS Homo sapiens.
FT Key Location/Qualifiers
FT misc.difference 29
FT /*tag- a
FT /*label- pyrimidine
PN W09421118-A.
PD 29-SEP-1994.
PE 24-MAR-1994; U03197.
PR 24-MAR-1993; US-035435.
PA (UABR-) UAB RES FOUND.
PI Garver RI, Sorscher EJ;
PI WPI: 94-316537/39.
DR DNA construct for treating human carcinoma - includes a
PT cancer-therapeutic gene under the control of a promoter and a gp.
PS Claim 1; Fig.6; 54pp; English.
CC This enhancer element is part of a DNA construct used for treating
CC human carcinoma which contains a cancer therapeutic protein under
CC the control of a promoter and 3 enhancer sequences in a specific 5'-3'
CC order. This enhancer element is derived from the flanking region of
CC the human epithelial cell cytokerin-8 gene.
SQ Sequence 32 BP; 7 A; 1 C; 8 G; 15 T;

Query Match 0.1%; Score 28; DB 1; Length 32;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18835 ACCCATCTCTACTATAAATACAAAT 18862
DB 28 ACCCATCTCTACTATAAATACAAAT 1

RESULT 11
ID 05168 standard; DNA: 40 BP.
AC 05168;
DE 21-JUL-1994 (first entry)
DE Sequence of primer for PCR amplification of HIV-LP Pt.1 isolate
KW Human immunodeficiency virus; HIV-LP; PCR primer; ss.
OS Synthetic.
PN W09400562-A.
PD 06-JAN-1994.
PE 23-JUN-1993; U06162.
PR 24-JUN-1993; US-803421.
PA (CORR) CORRELL RES FOUND INC.
PA (OTRI) UNIV NEW YORK MT SINAI.
PA SCHOOL MEDICINE.
PI Gelman IH, Laurence JC;
PI WPI: 94-026200/03.
PT HIV-LP useful in vaccine formulations - is novel HIV virus
PT distinct from HIV-1 and or HIV-2 viruses
PS Example; page 7; 75pp; English.
CC HIV-LP is a new variant of the HIV family. A cDNA first strand was
CC synthesised from Pt. 1 pellet using MLV RT. The product was
CC converted into dsDNA and this cDNA was amplified by PCR using
CC primers 05167 and 05168.
SQ Sequence 40 BP; 3 A; 2 C; 3 G; 32 T;

Query Match 0.1%; Score 28; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4696 AAAAAAAAAAAAAAAAAAGAAAT 4723
DB 34 AAAAAAAAAAAAAAAAAAGAAAT 7

RESULT 12
ID T04081 standard; DNA: 48 BP.
AC T04081;
DE 16-MAY-1996 (first entry)
DE Trypsin inhibitor protein cDNA antisense primer-1.
KW Trypsin; inhibitor; human 198g cells; pancreatitis; shock; DIC;
KW multiple organ failure; disseminated intravascular coagulation; ss.
OS Synthetic.
PN J07242700-A.
PD 19-SEP-1995.
PE 04-MAR-1994; 059906.
PR 04-MAR-1994; JP-059906.
PA (MOCH) MOCHIDA PHARM CO LTD.
PA WPI: 95-355285/46.
DR Polypeptide having trypsin inhibitory activity - for the treatment
PT of pancreatitis, shock, multi-organ failure, etc.
PS Example 7; Page 15; 25pp; Japanese.
CC A polypeptide having trypsin inhibitory activity, mol.wt. 22-28 KD
CC or 17-23 KD (as determined by SDS-PAGE under reducing or non-
CC reducing conditions, respectively) and an N-terminal sequence as in
CC R/9191 is claimed. The polypeptide was isolated from human 198g
CC cells and is useful for treating pancreatitis, shock, multiple
CC organ failure and disseminated intravascular coagulation. The
CC present sequence is that of an antisense primer used for cloning
CC 198g cDNA coding for the full-length polypeptide.
SQ Sequence 48 BP; 2 A; 1 C; 1 G; 44 T;

Query Match 0.1%; Score 28; DB 1; Length 48;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4696 AAAAAAAAAAAAAAAAAAGCAT 4723
DB 32 AAAAAAAAAAAAAAAAAAGCAT 5

RESULT 13
V02146 standard; CDNA; 49 BP.
AC V02146;
DE 12-MAY-1998 (first entry)
KW Human secreted protein AK533 3' portion including the polyA tail.
KW Human secreted protein; ATCC 98026; cytokine; immunomodulation;
OS Homo sapiens.
PN WO9739123-A2.
PD 23-OCT-1997.
PF 14-APR-1997; U06139.
PR 18-APR-1996; US-634325.
PA (GENM) GENETICS INST INC.
PI Jacobs K, Lavallee ER, McCoy JM, Merberg D, Racie LA, Spaulding V, WPI; 97-52660/48.
DR New secreted proteins encoded clones present in ATCC 98026 - possibly having cytokine, cell proliferation/differentiation regulating, immunomodulating and many other activities
PS Disclosure; Page 86; 139pp; English.
CC The present sequence encodes a portion of a novel human secreted protein deposited under accession number ATCC 98026. The secreted protein can be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions and as nutritional supplements. It may also have a very wide range of biological activities although no evidence for any is provided in the specification. Typical of these are cytokine, cell proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human immunodeficiency virus infection, cancer, autoimmune diseases and allergy); regulation of hematopoiesis (e.g. for treating anaemia or as adjunct to chemotherapy); stimulation of growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for treating wounds, periodontal disease, neurological diseases stroke, fibrosis); inhibition or stimulation of follicle stimulating hormone (for control of fertility); chemotactic and chemokinetic activities (e.g. for treating infections, tumours); haemostatic or thrombolytic activity (e.g. for treating haemophilia, cardiac infarction etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as antimicrobials; for treating postlasis or other hyperproliferative disease; for regulation of metabolism, behaviour, and many others. Also contemplated is the use of the corresponding nucleic acid in gene therapy procedures.
Sequence 49 BP; 40 A; 2 C; 6 G; 1 T;

Query Match 0.1%; Score 28; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26678 GAAAAAAAAAAAAAAAAA 26705
DB 3 GAAAAAAAAAAAAAAAAA 30

RESULT 14
T88080 standard; CDNA; 49 BP.
AC T88080;
DE 14-MAY-1998 (first entry)
PT 3' portion of cDNA clone encoding secreted protein AK533.

KW Human; secreted protein; research; treatment; AK533;
KW 3' portion; ds.
OS Homo sapiens.
PN WO9739122-A2.
PD 23-OCT-1997.
PF 11-APR-1997; U06042.
PR 12-APR-1996; US-631184.
PA (MORO-) MORO PHARM INC.
PI Thecharides TC; WPI; 97-526459/48.
DR Human and murine secreted proteins - useful to research or treat diseases or disorders related to their function
PS Disclosure; Page 86; 140pp; English.
CC The present sequence is the 3' portion of a cDNA clone encoding a human secreted protein, which may have nutritional uses, or cytokine and cell proliferation/differentiation, immune stimulating or suppressing, haematopoiesis regulating, tissue growth, activin/inhibin, chemotactic/chemokinetic, haemostatic and thrombolytic, receptor/ligand, anti-inflammatory or tumour inhibition activities. It can also be used to research or treat diseases/disorders related to its function.
CC The partial cDNA clone AP162 was 1st isolated from a human adult placenta cDNA library. The partial cDNA clones AM931, AM610, AM340, AM282, AK647, AK583, AK533 and AK296 were 1st isolated from a human foetal kidney cDNA library. The partial cDNA clones H617 and B89 were 1st isolated from a human peripheral blood monocyte cell (Th1 or Th2) cDNA library. The partial cDNA clone AW191 was 1st isolated from a human ovary (PA-1 teratocarcinoma) cDNA library. The partial cDNA clones AT211, AT705 and AT319 were 1st isolated from a human lymphocyte and dendritic cell cDNA library. The partial cDNA clones AS34 and AS32 were 1st isolated from a human foetal brain cDNA library. The partial cDNA clone AR260 was 1st isolated from a human adult retina cDNA library. The partial cDNA clones K640 and K39 were 1st isolated from a murine bone marrow (stromal cell line FCW-4) cDNA library.
Sequence 49 BP; 40 A; 2 C; 6 G; 1 T;

Query Match 0.1%; Score 28; DB 1; Length 49;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26678 GAAAAAAAAAAAAAAAAA 26705
DB 3 GAAAAAAAAAAAAAAAAA 30

RESULT 15
V54636/c
ID V54636 standard; DNA; 53 BP.
AC V54636;
DE 18-NOV-1998 (first entry)
PT DNA detector sequence used in a telomeric activity assay.
KW Telomerase; telomeric repeat sequence; carcinogen; senescence; cancer; ss.
OS Synthetic.
FH Key
FT misc_binding
FT Location/Qualifiers
FT 31..54
FT /tag= a
FT /note= "forms a double-stranded region with bases 38-60 of V54639".
PN WO9837241-A1.
PD 27-AUG-1998.
PF 23-FEB-1998; U03725.
PR 18-JUN-1997; US-050109.
PR 24-FEB-1997; US-038798.
PA (MTE-) TM TECHNOLOGIES INC.
PI Benight AS, Faldasz BD, Lane MJ; WPI; 98-467588/40.
DR Detecting telomerase activity - without the need for polymerase chain reaction amplification, can be adapted to assay for telomerase inhibitors
PT Example 1; Fig 1C; 9pp; English.

Query Match	0.1%;	Score 28;	DB 1;	Length 53;
Best Local Similarity	100.0%;	Pred. No. 42;		
Matches	28;	Conservative	0;	Mismatches
				Indels

Qy	26678	GGAAAAAAAAAAAAAAAAAAAAAAAAA	26705
Db	32	GGAAAAAAAAAAAAAAAAAAAAAAAAA	5

Search completed: March 31, 2000, 10:14:26
Job time: 195864 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 23:51:03 ; Search time 339.96 Seconds
(without alignments)
10565.402 Million cell updates/sec

Title: US-08-852-495c-1_COPY_140000_170000

Perfect score: 30001
Sequence: 1 CCCCTCCATCCCATATGCAC.....TACATATTTATTTATGTA 30001

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 214294 segs, 59861574 residues

Word size : 0

Total number of hits satisfying chosen parameters: 349160

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database :
1: /cgn2_6/pdata1/1/lna/5A_COMB.seq:*
2: /cgn2_6/pdata1/1/lna/5B_COMB.seq:*
3: /cgn2_6/pdata1/1/lna/5C_COMB.seq:*
4: /cgn2_6/pdata1/1/lna/5D_COMB.seq:*
5: /cgn2_6/pdata1/1/lna/6_COMB.seq:*
6: /cgn2_6/pdata1/1/lna/PCRTUS_COMB.seq:*
7: /cgn2_6/pdata1/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	0.1	60	3	US-08-454-557C-57 Sequence 57, Appl
2	38	0.1	60	4	US-08-340-426D-57 Sequence 57, Appl
3	38	0.1	60	4	US-08-450-673C-57 Sequence 57, Appl
4	38	0.1	60	6	PCT-US95-17111A-57 Sequence 57, Appl
5	38	0.1	76	3	US-08-454-557C-69 Sequence 69, Appl
6	38	0.1	76	4	US-08-340-426D-69 Sequence 69, Appl
7	38	0.1	76	4	US-08-450-673C-69 Sequence 69, Appl
8	38	0.1	76	6	PCT-US95-17111A-69 Sequence 69, Appl
9	29	0.1	30	3	US-08-454-557C-6 Sequence 6, Appl
10	29	0.1	30	4	US-08-340-426D-6 Sequence 6, Appl
11	29	0.1	30	4	US-08-450-673C-6 Sequence 6, Appl
12	29	0.1	30	6	PCT-US95-17111A-6 Sequence 6, Appl
13	29	0.1	55	4	US-08-778-494B-67 Sequence 109, Appl
14	29	0.1	58	4	US-08-778-494B-109 Sequence 91, Appl
15	29	0.1	84	3	US-08-454-557C-91 Sequence 91, Appl
16	29	0.1	84	4	US-08-340-426D-91 Sequence 91, Appl
17	29	0.1	84	4	US-08-450-673C-91 Sequence 91, Appl
18	29	0.1	84	6	PCT-US95-17111A-91 Sequence 91, Appl
19	28	0.1	50	1	US-08-233-609-5 Sequence 5, Appl
20	28	0.1	50	1	US-08-444-083-5 Sequence 5, Appl
21	28	0.1	50	1	US-08-286-304-5 Sequence 5, Appl
22	28	0.1	50	1	US-08-442-745-5 Sequence 5, Appl
23	28	0.1	50	1	US-08-443-129-5 Sequence 5, Appl
24	28	0.1	50	1	US-08-443-952-5 Sequence 5, Appl
25	28	0.1	50	2	US-08-443-130-5 Sequence 5, Appl
26	28	0.1	50	6	PCT-US95-04467-5 Sequence 5, Appl

27	28	0.1	60	3	US-08-454-557C-58 Sequence 58, Appl
28	28	0.1	60	3	US-08-454-557C-58 Sequence 58, Appl
29	28	0.1	60	4	US-08-340-426D-58 Sequence 58, Appl
30	28	0.1	60	4	US-08-340-426D-58 Sequence 58, Appl
31	28	0.1	60	4	US-08-450-673C-58 Sequence 58, Appl
32	28	0.1	60	4	US-08-450-673C-58 Sequence 58, Appl
33	28	0.1	60	6	PCT-US95-17111A-58 Sequence 58, Appl
34	28	0.1	60	6	PCT-US95-17111A-58 Sequence 58, Appl
35	28	0.1	78	3	US-08-454-557C-70 Sequence 70, Appl
36	28	0.1	78	3	US-08-454-557C-70 Sequence 70, Appl
37	28	0.1	78	4	US-08-340-426D-70 Sequence 70, Appl
38	28	0.1	78	4	US-08-340-426D-70 Sequence 70, Appl
39	28	0.1	78	4	US-08-450-673C-70 Sequence 70, Appl
40	28	0.1	78	4	US-08-450-673C-70 Sequence 70, Appl
41	28	0.1	78	6	PCT-US95-17111A-70 Sequence 70, Appl
42	28	0.1	78	6	PCT-US95-17111A-70 Sequence 70, Appl
43	27	0.1	33	7	5478746-1 Patent No. 5478746
44	27	0.1	42	1	US-07-875-167-2 Sequence 2, Appl
45	27	0.1	42	1	US-08-287-164-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-454-557C-57
Sequence 57, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454, 557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36, 203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-57
Query Match 0.1%; Score 38; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 9496 GCTATTTTGTATTTTGTATTTAGACAGACGGTTTCACC 9533
Db 11 GCTATTTTGTATTTTGTATTTAGACAGACGGTTTCACC 48

Query Match 0.1%; Score 38; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9496 GCTAATTTTGTATTAGTAGACAGGCGTTTCACC 9533
|||||
DB 11 GCTAATTTTGTATTAGTAGACAGGCGTTTCACC 48

RESULT 5

US-08-454-557C-69
; Sequence 69, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-454-557C-69

Query Match 0.1%; Score 38; DB 3; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9496 GCTAATTTTGTATTAGTAGACAGGCGTTTCACC 9533
|||||
DB 11 GCTAATTTTGTATTAGTAGACAGGCGTTTCACC 48

RESULT 6

US-08-340-426D-69
; Sequence 69, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-340-426D-69

Query Match 0.1%; Score 38; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9496 GCTAATTTTGTATTAGTAGACAGGCGTTTCACC 9533
|||||
DB 11 GCTAATTTTGTATTAGTAGACAGGCGTTTCACC 48

RESULT 7

US-08-450-673C-69
; Sequence 69, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540

FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-6

Query Match 0.1%; Score 29; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28388 TTCAGCGATTCTCTGCTCAGCTCC 28416
DB 30 TTCAGCGATTCTCTGCTCAGCTCC 2

RESULT 11
US-08-450-673C-6/C
Sequence 6, Application US/08450673C
Patent No. 594888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-6

Query Match 0.1%; Score 29; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28388 TTCAGCGATTCTCTGCTCAGCTCC 28416
DB 30 TTCAGCGATTCTCTGCTCAGCTCC 2

RESULT 12
PCT-US95-17111A-6/C
Sequence 6, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-6

Query Match 0.1%; Score 29; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28388 TTCAGCGATTCTCTGCTCAGCTCC 28416
DB 30 TTCAGCGATTCTCTGCTCAGCTCC 2

RESULT 13
US-08-778-494B-67/C
Sequence 67, Application US/08778494B
Patent No. 5962272
GENERAL INFORMATION:
APPLICANT: Chenchik, Alex
ATTORNEY/AGENT INFORMATION:
NAME: Diachenko, Luda
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-778-494B-67/C

CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/778,494B
FILING DATE: 03-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/582,562
FILING DATE: 03-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Doran R.
REGISTRATION NUMBER: 38,261
REFERENCE/DOCKET NUMBER: CL-7C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 55 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
US-08-778-494B-67

Query Match 0.1%; Score 29; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26680 AAAAAAAAAAAAAAAAAAGTA 26708
DB 51 AAAAAAAAAAAAAAAAAAGTA 23

RESULT 14
US-08-778-494B-109/C
Sequence 109, Application US/0878494B
Patent No. 5962272
GENERAL INFORMATION:
APPLICANT: Chenchik, Alex
APPLICANT: Zhu, York
APPLICANT: Diachenko, Luda
APPLICANT: Siebert, Paul
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR FULL-LENGTH CDNA
TITLE OF INVENTION: CLONING
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/778,494B
FILING DATE: 03-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/582,562
FILING DATE: 03-JAN-1996

ATTORNEY/AGENT INFORMATION:
NAME: Pace, Doran R.
REGISTRATION NUMBER: 38,261
REFERENCE/DOCKET NUMBER: CL-7C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (352) 375-8100
TELEFAX: (352) 372-5800
INFORMATION FOR SEQ ID NO: 109:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (synthetic)
US-08-778-494B-109

Query Match 0.1%; Score 29; DB 4; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26680 AAAAAAAAAAAAAAAAAAGTA 26708
DB 52 AAAAAAAAAAAAAAAAAAGTA 24

RESULT 15
US-08-454-557C-91/C
Sequence 91, Application US/0845457C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Mandis, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3924
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-91

Query Match 0.1%; Score 29; DB 3; Length 84;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24038 TGTATCCAGCAGCTTTGGAGGCTGAG 24066

Mon Apr 3 08:24:08 2000

us-08-852-495c-1_copy_140000_170000.rni

Db 77 TGTATCCAGCAGCTTTGGAGGCTGAGG 49

Search completed: March 31, 2000, 10:04:17
Job time: 195319 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 30, 2000, 19:57:22 ; Search time 8179.71 Seconds
(without alignments)
13848.131 Million cell updates/sec

Title: us-08-852-495c-1_COPY_140000_170000
Perfect score: 30001
Sequence: 1 CCCGCCAATCCCATATGCAC.....TACATATTATATTATGTA 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 188781982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *

44: gb_est25: *
45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_gss1: *
80: gb_gss2: *
81: gb_gss3: *
82: gb_gss4: *
83: em_gss1: *
84: em_gss2: *
85: em_gss3: *
86: em_gss4: *
87: gb_gss5: *
88: gb_gss6: *
89: gb_gss7: *
90: gb_gss8: *
91: gb_gss9: *
92: em_gss5: *
93: em_gss6: *
94: em_gss7: *
95: em_gss8: *
96: em_gss9: *
97: em_gss10: *
98: em_gss11: *
99: gb_gss10: *
100: gb_gss11: *
101: em_gss12: *
102: gb_gss12: *
103: gb_gss13: *
104: gb_gss14: *
105: gb_gss15: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	0.1	97	25	N47846 yw95g08.r1

```

c 2 45 0.1 102 81 B62983
c 3 43 0.1 102 81 B48088
c 4 41 0.1 61 39 AA873656
c 5 41 0.1 95 20 T52775
c 6 40 0.1 72 37 AA715942
c 7 40 0.1 105 21 T63772
c 8 39 0.1 97 39 AA837701
c 9 39 0.1 102 81 B80126
c 10 38 0.1 104 27 C05051
c 11 37 0.1 80 50 F24490
c 12 37 0.1 96 28 AA078527
c 13 37 0.1 100 30 AA252633
c 14 37 0.1 101 25 N80946
c 15 36 0.1 74 28 N80946
c 16 36 0.1 87 20 D11650
c 17 36 0.1 95 40 AA923035
c 18 36 0.1 102 25 N83386
c 19 35 0.1 92 37 AA729064
c 20 35 0.1 95 35 AA578401
c 21 35 0.1 102 23 H39052
c 22 35 0.1 105 61 A1832832
c 23 35 0.1 105 99 A0196366
c 24 34 0.1 83 40 AA988425
c 25 34 0.1 87 20 D20989
c 26 34 0.1 87 21 T62174
c 27 34 0.1 37 22 R70733
c 28 33 0.1 77 62 A1914818
c 29 33 0.1 86 38 AA745546
c 30 33 0.1 100 34 AA515337
c 31 33 0.1 101 35 AA583697
c 32 33 0.1 102 81 B62983
c 33 33 0.1 104 39 AA831816
c 34 33 0.1 105 51 A1752407
c 35 32 0.1 80 31 AA289455
c 36 32 0.1 86 38 AA745546
c 37 32 0.1 102 36 AA654896
c 38 32 0.1 104 38 AA807731
c 39 32 0.1 105 81 B74378
c 40 31 0.1 53 43 A1206968
c 41 31 0.1 53 43 A1242004
c 42 31 0.1 53 44 A1249836
c 43 31 0.1 53 44 A1252536
c 44 31 0.1 54 34 AA457759
c 45 31 0.1 55 23 H51371

```

ALIGNMENTS

```

RESULT 1
LOCUS N47846
DEFINITION y959g8.r1 Soares placentae_8to9weeks_2NBHP8to9w Homo sapiens cDNA
clone IMAGE:260030 5' similar to contains Alu repetitive element;
mRNA sequence.
N47846
N47846.1 GI:1189012
EST
14-FEB-1996

```

```

ACCESSION N47846
VERSION N47846
KEYWORDS EST
SOURCE EST
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 97)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, B., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P., and
Wilson, R.
The Mashu-Merck EST Project
Unpublished (1995)
On May 9, 1995 this sequence version replaced gi:802689.
Contact: Wilson RK
Washington University School of Medicine

```

```

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNC; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Putative full length read
Seq primer: 17
High quality sequence stop: 129.
Location/Qualifiers
1..97
/organism="Homo sapiens"
/db_xref="GDB:3889736"
/db_xref="taxon:9606"
/clone="IMAGE:260030"
/cloae="Soares placentae_8to9weeks_2NBHP8to9w"
/der_stage="two placentae: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: placenta; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCTGACGAGGCGCGCATTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M. Fatima Bonaldi."

```

```

BASE COUNT 23 a 22 c 32 g 20 t
ORIGIN

```

```

Query Match 0.1%; Score 45; DB 25; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 19136 CACCACGCTCACTCCAGCTGGCGCAGAGGAGACTGTCTCA 19180
Db 53 CACCACGCTCACTCCAGCTGGCGCAGAGGAGACTGTCTCA 97

```

```

RESULT 2
LOCUS B62983
DEFINITION B62983 102 bp DNA GSS 21-JUN-1998
CIT-HSP-669F15.TP CIT-HSP Homo sapiens genomic clone 669F15,
genomic survey sequence.
B62983
B62983.1 GI:2636892
GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 102)
Kim, U.-J., Adams, K.D. and Simon, M.I.
Determination of clone end sequences of human Bacterial Artificial
Chromosomes
Unpublished (1997)
Contact: Ung-Jin Kim
Caltech Genome Research Lab
California Institute of Technology
Division of Biology, MS 147-75, Pasadena, CA 91125, USA
Tel: 626 796 7066
Fax: 626 395 4901
Email: ung@ash.tree.caltech.edu
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..102
/organism="Homo sapiens"

```

```

FEATURES
source

```

/db_xref="GDB:5491647"
/db_xref="taxon:9606"
/clone_1lb="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC1; site_1: HindIII; site_2: HindIII"

BASE COUNT 30 a 30 c 21 g 21 t

ORIGIN

Query Match 0.1%; Score 45; DB 81; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.00093; Indels 0; Gaps 0;
Matches 45; Conservative 0; Mismatches 0;

Oy 28457 GCTAATTTTGTATTTAGTAGAGATGGGTTTCACCATGTTGG 28501
|||||
Db 64 GCTAATTTTGTATTTAGTAGAGATGGGTTTCACCATGTTGG 20

RESULT 3
B48088/c 102 bp DNA GSS 08-APR-1999
LOCUS RPCI11-4N6.TP RPCI-11 Homo sapiens genomic clone RPCI-11-4N6,
DEFINITION genomic survey sequence.
ACCESSION B48088
VERSION B48088.1 GI:2600325
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 102)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
TITLE Use of BAC End Sequences for Sequence-Ready Map Building
JOURNAL unpublished (1997)
COMMENT Other_GSSs: RPCI11-4N6.TV
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@telong.med.buffalo.edu). Clones may be purchased from
BACPRC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

FEATURES
source location/Qualifiers
1..102
/organism="Homo sapiens"
/db_xref="GDB:7501469"
/db_xref="taxon:9606"
/clone_1lb="RPCI-11-4N6"
/clone_1lb="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; site_1: EcoRI; site_2: EcoRI;
RPCI11 Human Male BAC library"

BASE COUNT 32 a 24 c 25 g 21 t

ORIGIN

Query Match 0.1%; Score 43; DB 81; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 9501 TTTTGTATTTTGTAGTAGAGACAGGGTTTCACCATGTTGGCA 9543
|||||
Db 66 TTTTGTATTTTGTAGTAGAGACAGGGTTTCACCATGTTGGCA 24

RESULT 4
AA873656 61 bp mRNA EST 07-APR-1998
LOCUS CE02607.s1 NCI-CGAP_Ov2 Homo sapiens cDNA clone IMAGE:1384764
DEFINITION similar to contains Alu repetitive element; mRNA sequence.
ACCESSION AA873656
VERSION AA873656.1 GI:2969778
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 61)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL unpublished (1997)
COMMENT On May 8, 1995 this sequence version replaced gi:801221.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.C.E. Consortium/LINt at:
www-bio.11n1.gov/bdrp/image/image.html

Insert Length: 553 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 51.

FEATURES
source location/Qualifiers
1..61
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="IMAGE:1384764"
/clone_1lb="NCI-CGAP_Ov2"
/sex="female"
/tissue_type="ovary"
/lab_host="DH10B"
/note="Vector: PAMP10; mRNA made from invasive ovarian
tumor. cDNA made by oligo-dT priming. Non-directionally
cloned. Size-selected on agarose gel. average insert size
600 bp. Reference: Krizman et al. (1996) Cancer Research
56:5380-5383."

BASE COUNT 9 a 15 c 20 g 17 t

ORIGIN

Query Match 0.1%; Score 41; DB 39; Length 61;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 19122 GTGACCGAGATCGACACTGCACCTCAGCGTGGGACCA 19162
|||||
Db 61 GTGACCGAGATCGACACTGCACCTCAGCGTGGGACCA 21

RESULT 5
T52775/c 95 bp mRNA EST 06-FEB-1995
LOCUS Y27908.r1 Strategene ovary (#937217) Homo sapiens cDNA clone
DEFINITION IMAGE:67887 5' similar to contains Alu repetitive element; mRNA
sequence.
ACCESSION T52775

VERSION T52775.1 GI:654635
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 95)
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,
Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Roifling, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Other ESTs: ya79d08.s1
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Insert Size: 1013
High quality sequence stops: 82 Source: IMAGE Consortium, LNL This
clone is available royalty-free through LNL; contact the IMAGE
Consortium (info@image.llnl.gov) for further information.
Insert Length: 1013 Std Error: 0.00
Seq primer: M13Rpl
High quality sequence stop: 82.
Location/Qualifiers
1..95
/organism="Homo sapiens"
/db_xref="GDB:504144"
/db_xref="taxon:9606"
/clone_1db="IMAGE:67887"
/clone_1lb="Stratagene ovary (#937217)"
/sex="female"
/dev_stage="49 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: ovary; Vector: Bluescript SK; Site: 1: EcoRI;
Site: 2: XhoI; Cloned unidirectionally. Primer: Oligo dt.
Total ovary tissue, normal, caucasian. Average insert
size: 0.8 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5'
GAATCGGCGAGAG 3' -3' adaptor sequence: 5'
CTGAGTGTGTGTGTGTGTGT 3' "

BASE COUNT 19 a 32 c 20 g 24 t
ORIGIN

Query Match 0.1%; Score 41; DB 20; Length 95;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 732 AGGAGATCACTGACCTGGAGCAGAGGTTGACAGTGAG 772
|||||
Db 84 AGGAGATCACTGACCTGGAGCAGAGGTTGACAGTGAG 44

RESULT 6
LOCUS AA715942 72 bp mRNA EST 29-DEC-1997
DEFINITION nv77f06.t1 NCI_CGAP_Br4 Homo sapiens CDNA clone IMAGE:1235843
similar to contains Alu repetitive element; mRNA sequence.
ACCESSION AA715942
VERSION AA715942.1 GI:2728216
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 72)
COMMENT

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On May 8, 1995 this sequence version replaced gi:800927.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Kristina A. Cole, M.D.,
Ph.D. student, Rodrigo F. Chuagui, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.llnl.gov/dbp/image/image.html

Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 59.
Location/Qualifiers
1..72
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1db="IMAGE:1235843"
/clone_1lb="NCI_CGAP_Br4"
/sex="female"
/tissue_type="normal ductal tissue"
/lab_host="DH10B"
/note="Organ: breast; Vector: pAMP10; mRNA made from
normal breast ductal tissue. cDNA made by oligo-dt
priming. Non-directionally cloned. Size-selected on
agarose gel, average insert size 600 bp. Reference:
Kitzman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT 15 a 22 c 20 g 15 t
ORIGIN

Query Match 0.1%; Score 40; DB 37; Length 72;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19138 CCAGTCACTCAGCCTGGGCGACAGGAGGACTCTGTC 19177
|||||
Db 71 CCAGTCACTCAGCCTGGGCGACAGGAGGAGGACTCTGTC 32

RESULT 7
LOCUS T63772 105 bp mRNA EST 17-FEB-1995
DEFINITION YC21f01.s1 Stratagene lung (#937210) Homo sapiens CDNA clone
IMAGE:81337 3' similar to contains Alu repetitive element; mRNA
sequence.
ACCESSION T63772
VERSION T63772.1 GI:667637
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 105)
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,
Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Roifling, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

Insert Size: 513

High quality sequence stops: 69 Source: IMAGE Consortium, LNL. This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.

Insert Length: 513 Std Error: 0.00

Seq primer: -21m13

High quality sequence stop: 69.

Location/Qualifiers

1. 105

/organism="Homo sapiens"

/db_xref="GDB:486954"

/db_xref="taxon:9606"

/clone_image="IMAGE:81337"

/clone_id="Stratagene lung (#937210)"

/sex="male"

/dev_stage="72 years"

/lab_host="SOLR cells (kanamycin resistant)"

/note="Organ: lung; Vector: paluescript SK; Site_1:

ECORI; Site_2: XhoI; Cloned unidirectionally. Primer:

Oligo dt. normal lung. Average insert size: 1.0 kb;

uni-zap xr vector; -5' adaptor sequence: 5' GAATTCGCACGAG

3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"

BASE COUNT 17 a 34 c 29 g 25 t

ORIGIN

Query Match 0.1%; Score 40; DB 21; Length 105;

Best Local Similarity 100.0%; Pred. No. 0.031;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28372 ACCTCCGCTCCCGGTTCAAGCATCTCTGCTCAGC 28411

DB 36 ACCTCCGCTCCCGGTTCAAGCATCTCTGCTCAGC 75

LOCUS

AA837701 97 bp mRNA EST 07-APR-1998

DEFINITION

0606C02.s1 NCI-CGAP_Ov2 Homo sapiens cDNA IMAGE:1385090

similar to contains element PTR5 repetitive element ; mRNA

sequence.

AA837701

AA837701.1 GI:2912900

EST.

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 97)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

On Jan 17, 1998 this sequence version replaced g1:1900439.

Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: David B. Kitzman, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/INL at:

www-bio.lnl.gov/bdrip/image/image.html

Insert Length: 451 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 50.

FEATURES

Location/Qualifiers

1. 97

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="IMAGE:1385090"

/clone_id="NCI-CGAP_Ov2"

/sex="female"

/tissue_type="ovary"

/lab_host="DH10B"

/note="Vector: pAMP10; mRNA made from invasive ovarian

tumor, cDNA made by oligo-dt priming. Non-directionally

cloned. Size-selected on agarose gel, average insert size

600 bp. Reference: Kitzman et al. (1996) Cancer Research

56:5380-5383."

BASE COUNT 13 a 24 c 29 g 31 t

ORIGIN

Query Match 0.1%; Score 39; DB 39; Length 97;

Best Local Similarity 100.0%; Pred. No. 0.066;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 734 GAGATCACTGAACCTGGAGGAGGAGGTTGCACTGAG 772

DB 97 GAGATCACTGAACCTGGAGGAGGAGGTTGCACTGAG 59

LOCUS

B80126 102 bp DNA GSS 24-OCT-1998

DEFINITION

CIT-HSP-2042J24.TR CIT-HSP Homo sapiens genomic clone 2042J24,

genomic survey sequence.

B80126

B80126.1 GI:2867149

GSS.

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 102)

Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,

Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,

Simon,M. and Venter,J.C.

Use of a random BAC End Sequence Database for Sequence-Ready Map

Building (1998)

Unpublished (1998)

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html

Seq primer: M13 Reverse

Class: BAC ends.

Location/Qualifiers

1. 102

/organism="Homo sapiens"

/db_xref="GDB:705355"

/db_xref="taxon:9606"

/clone_image="IMAGE:2042J24"

/clone_id="CIT-HSP"

/sex="male"

/cell_type="Sperm"

/note="Vector: pBelobAC1; Site_1: HindIII; Site_2:

HindIII"

BASE COUNT 20 a 32 c 19 g 31 t

ORIGIN

Query Match 0.1%; Score 39; DB 81; Length 102;
 Best Local Similarity 100.0%; Pred. No. 0.065;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24024 CAGTGCGTCACGGCTGTAATCCAGACATTGGAGGCT 24062

Db 64 CAGTGCGTCACGGCTGTAATCCAGACATTGGAGGCT 102

RESULT 10

C05051/c

LOCUS C05051 104 bp mRNA EST 30-JUL-1996
 DEFINITION C05051 Human heart cDNA (Ynakamura) Homo sapiens cDNA clone
 3NHC4544, mRNA sequence.

ACCESSION C05051
 VERSION C05051.1 GI:1468302

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 104)
 Tanaka, T., Ogiwara, A., Uchiyama, I., Takagi, T., Yazaki, Y. and
 Nakamura, Y.

REFERENCE

AUTHORS

Construction of a normalized directionally cloned cDNA library from
 adult heart and analysis of 3040 clones by partial sequencing

JOURNAL

MEDLINE

COMMENT

CONTACT: Yusuke Nakamura

Institute of Medical Science

University of Tokyo

4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan

Tel: 81-3-5449-5372

Fax: 81-3-5449-5433

Email: yusuke@ims.u-tokyo.ac.jp.

Location/Qualifiers

1..104

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="3NHC4544"

/clone_lib="Human heart cDNA (Ynakamura)"

/dev_stage="adult"

/note="Organ: heart; normalized directionally cloned cDNA

from adult heart"

BASE COUNT 24 a 26 c 37 g 17 t

ORIGIN

Query Match 0.1%; Score 38; DB 27; Length 104;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28378 GCCTCCCGGCTCAGACGATTCCTCCGCTCAGCCTCC 28415

Db 38 GCCTCCCGGCTCAGACGATTCCTCCGCTCAGCCTCC 1

RESULT 11

F24490

LOCUS F24490 80 bp mRNA EST 13-MAY-1999

DEFINITION HSPD10834 HM3 Homo sapiens cDNA clone s4000013A06, mRNA sequence.

ACCESSION F24490

VERSION F24490.1 GI:4810116

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 80)

Lanfranchi, G., Muraro, T., Caldera, F., Pacchioni, B., Pallavicini, A.,

Pandolfo, D., Toppo, S., Trevisan, S., Scarso, S. and Valle, G.

Identification of 4370 expressed sequence tags from a

LOCATION/Qualifiers

3'-end-specific cDNA library of human skeletal muscle by DNA
 sequencing and filter hybridization
 Genome Res. 6 (1), 35-42 (1996)

On Jun 5, 1998 this sequence version replaced gi:3188834.

Contact: Valle G.

CIRI Biotechnology Centre

University of Padua

Via Trieste 75, 35121 Padua, Italy

ABI Chromatograms and other information are available on WWW at

http://grup.bio.unipd.it.

Location/Qualifiers

1..80

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="s4000013A06"

/clone_lib="HM3"

/sex="female"

/tissue_type="pectoral muscle (after mastectomy)"

/note="Vector: pCDN41 (Invitrogen); Site: 1: BstXI;

Site: 2: NotI; The library was constructed by G.

Lanfranchi. This library is not subtracted nor normalized.

The first strand cDNA was primed with a biotinylated

0190-dT-NotI primer

(5'-biotin-AACCGGCTGAGAGCGCCGCTTTTCTTTTCTTTT-3'). The

ds cDNA was sonicated and size-selected in the range

350-550 bp. The 3' specific fragments were selected by

streptavidin coated magnetic beads, ligated to

non-palindromic BstXI adapters, NotI digested and

directionally cloned into BstXI-NotI cut pCDN41 vector."

BASE COUNT 18 a 21 c 19 g 22 t

ORIGIN

Query Match 0.1%; Score 37; DB 50; Length 80;
 Best Local Similarity 100.0%; Pred. No. 0.3;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13521 GTGGCTCATGCTGTATCCAGACACTTGGAGGCC 13557

Db 44 GTGGCTCATGCTGTATCCAGACACTTGGAGGCC 80

RESULT 12

AA078527/c

LOCUS AA078527 96 bp mRNA EST 24-SEP-1999

DEFINITION 7P05B09 Chromosome 7 Placental cDNA library Homo sapiens cDNA clone

7P05B09, mRNA sequence.

ACCESSION AA078527

VERSION AA078527.1 GI:1838015

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 96)

Touchman, J.W., Bouffard, G., Weintrub, L.A., Idol, J.R., Wang, L.,

Robbins, C.M., Nussbaum, J.C., Lovett, M. and Green, E.D.

2006 expressed-sequence tags derived from human chromosome

7-enriched cDNA libraries

Genome Res. 7 (3), 281-292 (1997)

On Sep 12, 1996 this sequence version replaced gi:1288977.

Contact: Eric D. Green

Genome Technology Branch

National Human Genome Research Institute/NIH

49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892

Tel: 3014020201

Fax: 3014024735

Email: egreen@hghri.nih.gov

Plate: 05 row: B column: 09

Seq primer: -21M3 (ABI).

Location/Qualifiers

Job time: 171429 sec

Query Match 0.1%; Score 37; DB 25; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24101 CCTGCCACATGGTGAACCCGCTCTCTACTAAAT 24137
 ||||||||||||||||||||||||||||||||||||
 Db 39 CCTGCCACATGGTGAACCCGCTCTCTACTAAAT 3

RESULT 15

AA078709

LOCUS

DEFINITION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

On Sep 12, 1996 this sequence version replaced g1:1293976.

Contact: Eric D. Green

Genome Technology Branch

National Human Genome Research Institute/NIH

49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892

Tel: 3014020201

Fax: 3014024735

Email: egreen@nhgri.nih.gov

Plate: 01 row: E column: 11

Seq primer: -21M13 (AB1).

Location/Qualifiers

1. .74

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="7T01E11"

/clone.lib="Chromosome 7 Thymus cDNA Library"

/sex="mixture of female and male"

/tissue_type="thymus"

/dev_stage="1 week old (male), 11 year old (female)"

/lab_host="E. coli strain DH5 alpha"

/note="Organ: thymus; Vector: PAMPI0; cDNA was generated from cytoplasmic RNA using a mixture of random DNA hexamers and oligo(dT). From this pool of cDNA, human chromosome 7-enriched cDNA was isolated by direct cDNA selection using chromosome 7 genomic DNA (cosmids). The resulting direct-selected cDNA was cloned into a plasmid vector using a non-directional uracil DNA glycosylase (UDG)-mediated cloning strategy."

BASE COUNT

ORIGIN

19 a

17 c

22 g

16 t

Query Match

Best Local Similarity

Matches

36; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY 13521 GTGGCTCATGCTGTATCCAGCACTTTGGAGGC 13556

|||||

Db 39 GTGGCTCATGCTGTATCCAGCACTTTGGAGGC 74

Search completed: March 31, 2000, 02:52:55

Mon Apr 3 08:24:08 2000

us-08-852-495c-1_copy_140000_170000.rst

Page 9

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2000, 06:36:34 ; Search time 13412.9 Seconds

(without alignments)
-6791.568 Million cell updates/sec

Title: US-08-852-495C-1_COPY_165000_195000

Perfect score: 30001
Sequence: 1 AGTCATATTTATAGAGTCC.....GCACTTTTCATTCTCTGCT 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 821193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_om:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl1:*
10: gb_pl2:*
11: gb_pr3:*
12: gb_ro:*
13: gb_sts:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hum3:*
39: em_hum4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*

44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	0.2	60	5	I31332
2	50	0.2	91	5	I31254
3	50	0.2	92	5	I31518
4	46	0.2	46	5	I31437
5	42	0.1	56	5	I31153
6	41	0.1	50	5	I31467
7	40	0.1	40	5	A68622
8	36	0.1	40	5	A68621
9	36	0.1	40	5	A68621
10	36	0.1	103	9	HUMALCE221
11	35	0.1	35	5	A25212
12	35	0.1	35	5	A25212
13	35	0.1	35	24	E09140
14	35	0.1	35	24	E09140
15	35	0.1	91	13	HUMUT8164A
16	35	0.1	102	13	G37854
17	34	0.1	63	13	HUMUT5302A
18	34	0.1	72	13	G31952
19	34	0.1	75	3	BTIGA45
20	34	0.1	80	9	HUMBRKFEAB
21	34	0.1	80	9	HUMBRKFEAB
22	34	0.1	88	3	OAUI5702
23	34	0.1	95	13	A0026531
24	34	0.1	95	13	A0026788
25	33	0.1	33	5	I31400
26	33	0.1	40	5	I31263
27	33	0.1	46	5	I31159
28	33	0.1	51	5	I31405
29	33	0.1	54	5	AR002288
30	33	0.1	54	5	AR053139
31	33	0.1	60	34	DRORS1
32	33	0.1	62	5	I07633
33	33	0.1	79	10	S73203
34	33	0.1	82	9	HUMRSCB
35	33	0.1	87	9	HUWL
36	33	0.1	87	12	MMVIMV24
37	33	0.1	87	13	G31698
38	33	0.1	90	12	MMDNDS22
39	33	0.1	91	12	MMD3NDS1
40	33	0.1	95	12	MMD4NDS2
41	33	0.1	95	34	AG2H79
42	33	0.1	97	3	SHPMCRE
43	33	0.1	97	9	HUMLDLRA2
44	33	0.1	97	9	HUMLDLRDT
45	33	0.1	98	34	AG2H815

ALIGNMENTS

RESULT 1
LOCUS I31332 60 bp DNA
DEFINITION Sequence 244 from patent US 5582979.
ACCESSION I31332
VERSION I31332.1 GI:1822123 PAT 06-FEB-1997

KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 60)
AUTHORS	Weber,J.T.
TITLE	Length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and method of using the same
JOURNAL	Patent: US 5582979-A 244 10-DEC-1996;
FEATURES	Location/Qualifiers
SOURCE	1..60
BASE COUNT	18 a 30 c 0 g 12 t
ORIGIN	"/organism="unknown"

[illegible][illegible]

```

Query Match          0.2%; Score 50; DB 5; Length 91;
Best Local Similarity 100.0%; Pred. No. 3,8e-13;
Matches    50; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

QY  28932 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACACACA 28981
      |||||
DB   30 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACACACA 79

```

RESULT	3				
LOCUS	131518				
DEFINITION	131518	92 bp	DNA	PAT	06-FEB-1997
ACCESSION	131518	Sequence 430	from patent US 5582979.		
VERSION	131518.1	GI:1822309			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
	Unclassified.				
REFERENCE	1 (bases 1 to 92)				
AUTHORS	Weber, J.L.				
TITLE	length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and method of using the same				
JOURNAL	Patent: US 5582979-A 430 10-DEC-1996;				
FEATURES	Location/Qualifiers				
SOURCE	1..92				

	21 a	43 c	1 g	27 t
BASE COUNT				
ORIGIN				

	Query Match	0.2%;	Score 50;	DB 5e;	Length 92;
	Best Local Similarity	100.0%	Pred. No. 38-13;		
	Matches 50;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	28932	CTCTCTCTCTCTCTCTCTCTCAACACACACACACACACACACA	28981		
Db	30	CTCTCTCTCTCTCTCTCTCTCAACACACACACACACACACACA	79		

RESULT	4			PAT	06-FEB-1997
Locus	I11437				
DEFINITION	I11437	46 bp	DNA		
ACCESSION	I11437	Sequence 349 from patent US 5582979.			
VERSION	I11437.1	G1:1822228			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 46)				
AUTHORS	Weber J.L.				
TITLE	Length polymorphisms in (dc-da).sub.n.(dg-dr).sub.n sequences and method of using the same				
JOURNAL	Patent: US 5582979-A 349 10-DEC-1996;				
FEATURES	Location/Qualifiers				
Source	1..46				
	/organism="unknown"				
BASE COUNT	12 a	0 c	0 g	34 t	
ORIGIN					

```
Query Match Similarity      0.29; Score 46; DB 5; Length 46;
Best Local Similarity      100.0%; Pred. NO. 4e-11;
Matches    46; Conservative 0; Mismatches    0; Indels    0; Gaps    0;
```

RESULT	5				
LOCUS	131153				
DEFINITION	Sequence 65 from patent US 5582979.				PAT
ACCESSION	131153				06-FEB-1997
VERSION	131153.1				
KEYWORDS	GI:1821944				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 56)				
TITLE	Weber J.L.				
JOURNAL	Length polymorphisms in (dc-da).sub.n.(dg-dr).sub.n sequences and				
FEATURES	method of using the same				
	Patent: US 5582979-A 65 10-DEC-1996;				
	Location/Qualifiers				
	1..56				
	/organism="unknown"				
BASE COUNT	23 a	28 c	0 g	5 t	
ORIGIN					

	Query Match	0.18;	Score 42;	DB 5;	Length 56;	
	Best Local Similarity	100.00;	Pred. No. 3,7e-09;			
	Matches 42;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
OY	28944	CTCTCTCTCTCACACACACACACACACACACACACA	28985			

Db 1 CTCCTCTCTCACACACACACACACACACACACACA 42

RESULT 6
LOCUS I31467 50 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 379 from patent US 5582979.
ACCESSION I31467
VERSION I31467.1 GI:1822258
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 50)
AUTHORS Weber,J.L.
TITLE Length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and method of using the same
JOURNAL Patent: US 5582979-A 379 10-DEC-1996;
FEATURES Location/Qualifiers
source 1..50 /organism="unknown"
BASE COUNT 20 a 25 c 0 g 5 t
ORIGIN

Query Match 0.1%; Score 41; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28945 TCCTCTCTCTCACACACACACACACACACACACA 28985
Db 1 TCCTCTCTCTCACACACACACACACACACACACACA 41

RESULT 7
LOCUS A68622 40 bp DNA PAT 06-MAY-1999
DEFINITION Sequence 2 from Patent WO9801573.
ACCESSION A68622
VERSION A68622.1 GI:4759649
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Resnick,M.A., Larionov,V.L., Koupriina,N.Y. and Perkins,E.L.
TITLE TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
JOURNAL Patent: WO 9801573-A 15-JAN-1998;
FEATURES Location/Qualifiers
source 1..40 /organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 9 a 8 c 19 g 4 t
ORIGIN

Query Match 0.1%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3376 CCGCTCCCGGGTTCAGCGATTCCTCCGCTCAGCCCTCC 3415
Db 40 CCGCTCCCGGGTTCAGCGATTCCTCCGCTCAGCCCTCC 1

RESULT 8
LOCUS A68621 40 bp DNA PAT 06-MAY-1999
DEFINITION Sequence 1 from Patent WO9801573.
ACCESSION A68621
VERSION A68621.1 GI:4759648
KEYWORDS

SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Resnick,M.A., Larionov,V.L., Koupriina,N.Y. and Perkins,E.L.
TITLE TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
JOURNAL Patent: WO 9801573-A 15-JAN-1998;
FEATURES Location/Qualifiers
source 1..40 /organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 7 a 12 c 13 g 8 t
ORIGIN

Query Match 0.1%; Score 36; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19729 TGGCTCAGCCCTGTAATCCACGACCTTTGGAGGCC 19764
Db 3 TGGCTCAGCCCTGTAATCCACGACCTTTGGAGGCC 38

RESULT 9
LOCUS A68621 40 bp DNA PAT 06-MAY-1999
DEFINITION Sequence 1 from Patent WO9801573.
ACCESSION A68621
VERSION A68621.1 GI:4759648
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Resnick,M.A., Larionov,V.L., Koupriina,N.Y. and Perkins,E.L.
TITLE TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
JOURNAL Patent: WO 9801573-A 15-JAN-1998;
FEATURES Location/Qualifiers
source 1..40 /organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 7 a 12 c 13 g 8 t
ORIGIN

Query Match 0.1%; Score 36; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21783 GCCTCCCAAGTCTGGGATTCAGGCGTGAGCCAC 21818
Db 37 GCCTCCCAAGTCTGGGATTCAGGCGTGAGCCAC 2

RESULT 10
LOCUS HMA1CE221 103 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS Simeitt,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press

AC E09140;
SV E09140.1
DT 07-OCT-1997 (Rel. 52, Created)
DT 08-OCT-1997 (Rel. 52, last updated, Version 1)
XX Synthetic DNA for Alu specific primer.
DE
XX JP 1995115999-A/3.
XX
XX unidentified
OC unclassified.
XX
XX [1]
RN 1-35
RP Andorea H.A., Van F.;
RA "DETECTING METHOD FOR DNA ARRANGEMENT VARIATION";
RT Patent number JP 1995115999-A/3, 09-MAY-1995.
RL INGENII BV.
XX
XX OS None
CC Artificial sequences.
CC PN JP 1995115999-A/3
CC PD 09-MAY-1995
CC PF 22-MAY-1992 JP 1992130668
CC PI ANOREASU HERARDODUSU AITSUTERURINDEN, YAN FUEIKU
CC PC C12Q1/68, C12N15/00, G01N27/447, G01N27/447;
CC CC strandedness: Single;
CC CC topology: Linear;
CC CC Key Location/Qualifiers
CC FH 1.35
CC FT /organism="Artificial sequences"
CC FT misc.feature 1.35
CC FT /note="Alu specific primer"
FH Key Location/Qualifiers
FH FT source 1.35
FH FT /db_xref="taxon:32644"
FH FT /organism="unidentified"
XX
XX SO Sequence 35 Bp: 8 A; 10 C; 11 G; 6 T; 0 other;

Query Match 0.1%; Score 35; DB 24; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19729 TGCGTACACGCTGTATCCGACACTTGGAGGC 19763
DB 35 TGCGTACACGCTGTATCCGACACTTGGAGGC 1

RESULT 15
LOCUS HUMUT8164A 91 bp DNA STS 29-DEC-1994
DEFINITION Human STS UT8164, 5' primer bind, sequence tagged site.
ACCESSION L30244
VERSION L30244.1 GI:605447
KEYWORDS STS; PCR primer; STS sequence; microsatellite DNA; microsatellite
marker; sequence tagged site; tetranucleotide repeat.
SOURCE Homo sapiens DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 91)
AUTHORS Gerken,S.C., Matsunami,N., Plaetke,R., Albertsen,E., Ballard,L.,
Wells,R., Lawrence,E., Moore,M., Holik,P.R., Carlson,M., Zhao,X.,
Robertson,M., Bradley,P., Elsner,T., Tinsley,A., Lalouel,J.-M. and
White,R.
TITLE Genetic and physical mapping of simple sequence repeat containing

JOURNAL Unpublished (1994)
COMMENT Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: stegcorona.med.utah.edu
Primer A: AGAGCTGCGACGACGACAA
Primer B: TTTTCCCCCTCTACTACT
End to Label: Primer B
PCR Profile:
Initial Denaturation: 94C 300sec
Cycles Denaturation Annealing Extension 5 94
C 10 sec. 56 C 10 sec. 72 C 20 sec. 30
52 C 10 sec. 72 C 20 sec. Mg++: 1.50 mM
Gel: Acrylamide 7%, Formamide 32%, Urea 348
Alleles: 1.

FEATURES Location/Qualifiers
source 1.91
/organism="Homo sapiens"
/db_xref="taxon:9606"
primer_bind 10.28
/evidence=experimental 13 t
BASE COUNT 35 a 20 c 23 g
ORIGIN

Query Match 0.1%; Score 35; DB 13; Length 91;
Best Local Similarity 100.0%; Pred. No. 9.9e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3317 GAGCTCAGCTGTGTCACCCGAGCTGGAGTGCAGTG 3351
DB 71 GAGCTCAGCTGTGTCACCCGAGCTGGAGTGCAGTG 37

Search completed: March 31, 2000, 17:25:40
Job time: 223289 sec

Sequence 88 BP; 56 A; 5 C; 7 G; 20 T; 5 Q

	Query Match	0.28;	Score 63;	DB 1;	Length 88;
	Best Local Similarity	100.0%;	Pred. No.	5.9e-08;	
	Matches 63; Conservative	0;	Mismatches	0;	Indels 0;
Qy_21514	TATTATTTTATTATTATTTTATTATTATTATTATTATTATTATTATTTTGAGACAGG	21573			

DE Alu PCR primer 1.
 KM PCR; primer; amplification; Alu repeat sequence; vector;
 KM circular yeast artificial chromosome; YAC; ss.
 OS Synthetic.
 OS Saccharomyces sp.
 PN WO9801573-A1.
 PD 15-JAN-1998.
 PF 09-JUL-1996; U11478.
 PR 09-JUL-1996; WO-U11478.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Koupina NY, Lariouov VL, Perkins EL, Resnick MA;
 PI WPI; 98-110234/10.
 PT Preparation of yeast artificial chromosomes - by in vivo
 PT recombination using vector comprising yeast centromere, marker,
 PT yeast telomere and nucleic acid for recombination
 PS Example 1; Page 45; 117pp; English.
 CC This is the nucleotide sequence for the PCR primer used in the
 CC amplification of the Alu repeat sequence, which is used to
 CC demonstrate the processes described in the invention. It involves
 CC the creation and use of circular yeast artificial chromosome (YAC)
 CC to selectively clone specific nucleic acids from a background of
 CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
 CC They can be used to rapidly isolate human DNA where only a part of the
 CC sequence of DNA is known. Using the methods large fragments of DNA can
 CC be easily cloned and analysed.
 SQ Sequence 40 BP; 7 A; 12 C; 13 G; 8 T;

Query Match 0.1%; Score 36; DB 1; Length 40;
 Best Local Similarity 100.0%; Pred.No.0.56;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19729 TGCGTCACGCGCTGTAATCCAGCACTTTGGAGGCC 19764
 DB 3 TGCGTCACGCGCTGTAATCCAGCACTTTGGAGGCC 38

RESULT 15
 VI9044/C
 ID VI9044 standard; DNA; 40 BP.
 AC VI9044;
 DT 28-JUL-1998 (first entry)
 DE Alu PCR primer 1.
 KM PCR; primer; amplification; Alu repeat sequence; vector;
 KM circular yeast artificial chromosome; YAC; ss.
 OS Synthetic.
 OS Saccharomyces sp.
 PN WO9801573-A1.
 PD 15-JAN-1998.
 PF 09-JUL-1996; U11478.
 PR 09-JUL-1996; WO-U11478.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Koupina NY, Lariouov VL, Perkins EL, Resnick MA;
 PI WPI; 98-110234/10.
 PT Preparation of yeast artificial chromosomes - by in vivo
 PT recombination using vector comprising yeast centromere, marker,
 PT yeast telomere and nucleic acid for recombination
 PS Example 1; Page 45; 117pp; English.
 CC This is the nucleotide sequence for the PCR primer used in the
 CC amplification of the Alu repeat sequence, which is used to
 CC demonstrate the processes described in the invention. It involves
 CC the creation and use of circular yeast artificial chromosome (YAC)
 CC to selectively clone specific nucleic acids from a background of
 CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
 CC They can be used to rapidly isolate human DNA where only a part of the
 CC sequence of DNA is known. Using the methods large fragments of DNA can
 CC be easily cloned and analysed.
 SQ Sequence 40 BP; 7 A; 12 C; 13 G; 8 T;

Query Match 0.1%; Score 36; DB 1; Length 40;
 Best Local Similarity 100.0%; Pred.No.0.56;
 Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21783 GCCTCCCAAGTGTGGATTACAGCGCTGAGCCAC 21818
 DB 37 GCCTCCCAAGTGTGGATTACAGCGCTGAGCCAC 2

Search completed: March 31, 2000, 19:25:47
 Job time: 228945 sec

Mon Apr 3 08:24:10 2000

us-08-852-495c-1_copy_165000_195000.png

Page 7

Mon Apr 3 08:24:11 2000

us-08-852-495c-1_copy_165000_195000.rml

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2000, 10:04:17 ; Search time 340.38 Seconds
(Without alignments)
10552.365 Million cell updates/sec

Title: US-08-852-495c-1_COPY_165000_195000

Perfect score: 30001
Sequence: 1 AGTGAAATTTATAGAGTCC.....GCAGTTTCATTGTGTGTGT 30001.

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 214294 seqs, 59861574 residues

Word size : 0

Total number of hits satisfying chosen parameters: 349160

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/lna/5C.COMB.seq.*
4: /cgn2_6/ptodata/1/lna/5D.COMB.seq.*
5: /cgn2_6/ptodata/1/lna/6.COMB.seq.*
6: /cgn2_6/ptodata/1/lna/PTUS9.COMB.seq.*
7: /cgn2_6/ptodata/1/lna/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	0.2	60	1	US-08-222-177A-244 Sequence 244, App
2	50	0.2	91	1	US-08-222-177A-166 Sequence 166, App
3	50	0.2	92	1	US-08-222-177A-430 Sequence 430, App
4	46	0.2	46	1	US-08-222-177A-349 Sequence 349, App
5	42	0.1	56	1	US-08-222-177A-65 Sequence 65, App1
6	41	0.1	50	1	US-08-222-177A-379 Sequence 379, App
7	33	0.1	33	1	US-08-222-177A-312 Sequence 312, App
8	33	0.1	40	1	US-08-222-177A-175 Sequence 175, App
9	33	0.1	46	1	US-08-222-177A-71 Sequence 71, App1
10	33	0.1	51	1	US-08-222-177A-317 Sequence 317, App
11	33	0.1	54	2	US-08-469-802B-27 Sequence 27, App1
12	33	0.1	54	3	US-08-267-803B-45 Sequence 45, App1
13	33	0.1	62	7	5506118-3 Patent No. 5506118
14	32	0.1	33	1	US-08-222-177A-59 Sequence 59, App1
15	32	0.1	33	1	US-08-222-177A-218 Sequence 218, App
16	32	0.1	34	1	US-08-222-177A-110 Sequence 110, App
17	32	0.1	34	1	US-08-222-177A-112 Sequence 112, App
18	32	0.1	34	1	US-08-222-177A-180 Sequence 180, App
19	32	0.1	34	1	US-08-222-177A-192 Sequence 192, App
20	32	0.1	34	1	US-08-222-177A-322 Sequence 322, App
21	32	0.1	35	1	US-08-222-177A-77 Sequence 77, App1
22	32	0.1	35	1	US-08-222-177A-189 Sequence 189, App
23	32	0.1	36	1	US-08-222-177A-104 Sequence 104, App
24	32	0.1	36	1	US-08-222-177A-358 Sequence 358, App
25	32	0.1	37	1	US-08-222-177A-140 Sequence 140, App
26	32	0.1	38	1	US-08-222-177A-198 Sequence 198, App

27	32	0.1	38	1	US-08-222-177A-397 Sequence 397, App
28	32	0.1	39	1	US-08-222-177A-137 Sequence 137, App
29	32	0.1	40	1	US-08-222-177A-119 Sequence 119, App
30	32	0.1	40	1	US-08-222-177A-400 Sequence 400, App
31	32	0.1	40	1	US-08-222-177A-403 Sequence 403, App
32	32	0.1	41	1	US-08-222-177A-74 Sequence 74, App1
33	32	0.1	41	1	US-08-222-177A-183 Sequence 183, App
34	32	0.1	41	1	US-08-222-177A-232 Sequence 232, App
35	32	0.1	42	1	US-08-222-177A-53 Sequence 53, App1
36	32	0.1	42	1	US-08-222-177A-229 Sequence 229, App
37	32	0.1	43	1	US-08-222-177A-370 Sequence 370, App
38	32	0.1	43	1	US-08-222-177A-388 Sequence 388, App
39	32	0.1	44	1	US-08-222-177A-195 Sequence 195, App
40	32	0.1	44	1	US-08-222-177A-241 Sequence 241, App
41	32	0.1	44	1	US-08-222-177A-249 Sequence 249, App
42	32	0.1	45	1	US-08-222-177A-157 Sequence 157, App
43	32	0.1	45	1	US-08-222-177A-203 Sequence 203, App
44	32	0.1	45	1	US-08-222-177A-364 Sequence 364, App
45	32	0.1	46	1	US-08-222-177A-101 Sequence 101, App

ALIGNMENTS

RESULT 1
US-08-222-177A-244
Sequence 244, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Demilt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 244:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd67rs
US-08-222-177A-244
Query Match 0.2%; Score 54; DB 1; Length 60;

APPLICATION NUMBER: US 07/341,562
 FILING DATE: 21-APR-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Sara, Charles S
 REGISTRATION NUMBER: 30,492
 REFERENCE/POCKET NUMBER: 03865,601
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (608) 831-2100
 TELEFAX: (608) 831-2106
 TELEX:
 INFORMATION FOR SEQ ID NO: 349:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 46 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 IMMEDIATE SOURCE:
 CLONE: mid1010rs
 US-08-222-177A-349

Query Match	0.28;	Score 46;	DB 1;	Length 46;
Best Local Similarity	100.0%;	Pred. No. 3.9e-06;		
Matches 46;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

```

1 RESULT 5
2 US-08-222-177A-65
3 : Sequence 65, Application US/08222177A
4 : Patent No. 5582979
5 :
6 : GENERAL INFORMATION:
7 :
8 : APPLICANT: Weber, James L.
9 :
10 : TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
11 : TITLE OF INVENTION: (dc-da)n.(dc-dt)n SEQUENCES AND METHODS OF USING SAME
12 : NUMBER OF SEQUENCES: 460
13 : CORRESPONDENCE ADDRESS:
14 : ADDRESSEE: Dewitt Ross & Stevens, S.C.
15 : STREET: 8000 Excelsior Drive, Suite 401
16 : CITY: Madison
17 : STATE: Wisconsin
18 : COUNTRY: USA
19 : ZIP: 53717-1914
20 :
21 : COMPUTER READABLE FORM:
22 : MEDIUM TYPE: Floppy disk
23 : COMPUTER: IBM PC compatible
24 : OPERATING SYSTEM: PC-DOS/MS-DOS
25 : SOFTWARE: Patentln Release #1.0, Version #1.25
26 :
27 : CURRENT APPLICATION DATA:
28 : APPLICATION NUMBER: US/08/222,177A
29 : FILING DATE:
30 : CLASSIFICATION: 435
31 :
32 : PRIOR APPLICATION DATA:
33 : APPLICATION NUMBER: US 07/341,562
34 : FILING DATE: 21-APR-1989
35 :
36 : ATTORNEY/AGENT INFORMATION:
37 : NAME: Sara, Charles S.
38 : REGISTRATION NUMBER: 30,492
39 : REFERENCE/DOCKET NUMBER: 09865,601
40 : TELECOMMUNICATION INFORMATION:
41 : TELEPHONE: (608) 831-2100
42 : TELEFAX: (608) 831-2106
43 :
44 : TELEX:
45 :
46 : INFORMATION FOR SEQ ID NO: 65:
47 :
48 : SEQUENCE CHARACTERISTICS:
49 : LENGTH: 56 base pairs
50 : TYPE: nucleic acid
51 : STRANDEDNESS: double
52 : TOPOLOGY: linear
53 :
54 :
55 :
56 :
57 :
58 :
59 :
60 :
61 :
62 :
63 :
64 :
65 :
66 :
67 :
68 :
69 :
70 :
71 :
72 :
73 :
74 :
75 :
76 :
77 :
78 :
79 :
80 :
81 :
82 :
83 :
84 :
85 :
86 :
87 :
88 :
89 :
90 :
91 :
92 :
93 :
94 :
95 :
96 :
97 :
98 :
99 :
100 :

```

```

; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: mfd5rs
US-08-222-177A-65

```

```
Query Match      0.1%; Score 42; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 28944 CTCTCTCTCTCACACACACACACACACACACA 28985
      |||||
Db 1 CTCTCTCTCTCACACACACACACACACACACA 42

```

RESULT 6
US-08-222-177A-379
Sequence 379, Application US/08222177A

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE F

```

1  COMPUTER READABLE FORM:
2
3  MEDIUM TYPE: Floppy disk
4
5  COMPUTER: IBM PC compatible
6
7  OPERATING SYSTEM: PC-DOS/MS-DOS
8
9  SOFTWARE: PatentIn Release #1.0, Version #1.25
10
11 CURRENT APPLICATION DATA:

```

```

; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865, 603
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106

```

? TELE:
? INFORMATION FOR SEQ ID NO: 379:
? SEQUENCE CHARACTERISTICS:
?     LENGTH: 50 base pairs
?     TYPE: nucleic acid
?     STRANDEDNESS: double
?     TOPOLOGY: linear
?     MOLECULE TYPE: DNA (genomic)

```

;
; IMMEDIATE SOURCE;
; CLONE: mfd120rs
; US-08-222-177A-379

Query Match	0.1%	Score 41:	DB 1:	Length 50:
Best Local Similarity	100.0%	Pred. No.	0.00015:	
Matches 41; Conservative	0;	Mismatches	0;	Indels 0
				Gaps 0

```
OY 28945 TCTCTCTCTCACACACACACACACACACACACACA 28985
      |||||
Db   1 TCTCTCTCTCACACACACACACACACACACACACA 41
```

RESULT 7

US-08-222-177A-312
Sequence 312, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865,601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 312:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd97rs
US-08-222-177A-312

Query Match 0.1%; Score 33; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.056;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28944 CTCCTCTCTCACACACACACACACAC 28976
|||||
DB 1 CTCCTCTCTCACACACACACACACACAC 33

RESULT 8
US-08-222-177A-175
Sequence 175, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865,601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 175:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd42rs
US-08-222-177A-175

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865,601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 175:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd42rs
US-08-222-177A-175

Query Match 0.1%; Score 33; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28954 CACACACACACACACACACACACACAC 28986
|||||
DB 1 CACACACACACACACACACACACACACAC 33

RESULT 9
US-08-222-177A-71
Sequence 71, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865,601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd42rs
US-08-222-177A-71

TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-222-177A-71

Query Match 0.1%; Score 33; DB 1; Length 46;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28954 CACACACACACACACACACACACACAT 28986
Db 9 CACACACACACACACACACACACACAT 41

RESULT 10
US-08-222-177A-317
Sequence 317, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 317:
SEQUENCE CHARACTERISTICS:
LENGTH: 51 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfg993c
US-08-222-177A-317

Query Match 0.1%; Score 33; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28953 TCACACACACACACACACACACACACA 28985
Db 10 TCACACACACACACACACACACACACA 42

RESULT 11
US-08-469-802B-27
Sequence 27, Application US/08469802B
Patent No. 5741645
GENERAL INFORMATION:
APPLICANT: Orr, Harry T.
APPLICANT: Raun, Laura P.W.
APPLICANT: Chung, Ming-Yi
APPLICANT: Zoghbi, Huda Y.
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
Patent No. 5741645
TITLE OF INVENTION: Type 1 and Method for Diagnosis
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mueling, Raasch, Gebhardt & Schwappach, P.A.
STREET: 119 No. 5741645th Fourth Street, Suite 203
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,802B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mueling, Ann M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 110,00030101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1225
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-469-802B-27

Query Match 0.1%; Score 33; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28954 CACACACACACACACACACACACACAT 28986
Db 1 CACACACACACACACACACACACACAT 33

RESULT 12
US-08-267-803B-45
Sequence 45, Application US/08267803B
Patent No. 5834183
GENERAL INFORMATION:
APPLICANT: Orr, Harry T.
APPLICANT: Raun, Laura P.W.
APPLICANT: Chung, Ming-Yi
APPLICANT: Zoghbi, Huda Y.
TITLE OF INVENTION: Gene Sequence for Spinocerebellar Ataxia
Patent No. 5834183
TITLE OF INVENTION: Type 1 and Method for Diagnosis

NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Mueling, Raasch, Gebhardt & Schwappach, P.A.
STREET: P.O. Box 581415
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55458-1415
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/267, 803B
FILING DATE: 28-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McCormack, Myra H.
REGISTRATION NUMBER: 36,602
REFERENCE/DOCKET NUMBER: 110,00030120
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-267-803B-45

Query Match 0.1%; Score 33; DB 3; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28954 CACACACACACACACACACACACACACAT 28986
1 CACACACACACACACACACACACACACAT 33

RESULT 13
5506118-3/C
PATENT NO. 5506118
APPLICANT: BERG, DAVID T.; GRINNELL, BRIAN W.
TITLE OF INVENTION: METHOD OF USING EUKARYOTIC EXPRESSION
VECTORS COMPRISING A POLY GT ELEMENT IN THE PRESENCE OF
TRANS-ACTING GENE PRODUCTS
NUMBER OF SEQUENCES: 13
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/110,475
FILING DATE: 23-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 914,651
FILING DATE: 15-JUL-1992
APPLICATION NUMBER: 255,203
FILING DATE: 07-OCT-1988
SEQ ID NO: 3;
LENGTH: 62
5506118-3

Query Match 0.1%; Score 33; DB 7; Length 62;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28953 TCACACACACACACACACACACACACA 28985
53 TCACACACACACACACACACACACACACA 21

RESULT 14
US-08-222-177A-59
Sequence 59, Application US/08222177A
Patent No. 5582979

GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989

ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865, 601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106

INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd3rs

US-08-222-177A-59

Query Match 0.1%; Score 32; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28954 CACACACACACACACACACACACACACA 28985
1 CACACACACACACACACACACACACACA 32

RESULT 15
US-08-222-177A-218
Sequence 218, Application US/08222177A
Patent No. 5582979

GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2000, 02:52:55 ; Search time 8179.94 Seconds
(without alignments)
13847.741 Million cell updates/sec

Title: US-08-852-495c-1_COPY_165000_195000
Perfect score: 30001
Sequence: 1 AGTGAATTTATATAGATGC.....GCAGTTTTCATTGTCTGCT 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8
Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *

44: gb_est25: *
45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_gss1: *
80: gb_gss2: *
81: gb_gss3: *
82: gb_gss4: *
83: em_gss1: *
84: em_gss2: *
85: em_gss3: *
86: em_gss4: *
87: gb_gss5: *
88: gb_gss6: *
89: gb_gss7: *
90: gb_gss8: *
91: gb_gss9: *
92: em_gss5: *
93: em_gss6: *
94: em_gss7: *
95: em_gss8: *
96: em_gss9: *
97: em_gss10: *
98: em_gss11: *
99: gb_gss10: *
100: gb_gss11: *
101: em_gss12: *
102: gb_gss12: *
103: gb_gss13: *
104: gb_gss14: *
105: gb_gss15: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	45	0.1	75	25	N68192

N68192 za11f07.s1

/db_xref="taxon:9606"
/clone="669F15"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT 30 a 30 c 21 g 21 t
ORIGIN

Query Match 0.1%; Score 45; DB 81; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3457 GCTAATTTTGTATTTTGTAGAGATGGGTTTCACCATGTTGG 3501
|||||
Db 64 GCTAATTTTGTATTTTGTAGAGATGGGTTTCACCATGTTGG 20

RESULT 3
AA053038 91 bp mRNA EST 13-SEP-1996
LOCUS
DEFINITION 2171612.s1 StrataGene colon (#937204) Homo sapiens CDNA clone
IMAGE:510094 3' similar to contains Alu repetitive element; mRNA
sequence.
ACCESSION AA053038 GI:1544176
VERSION
KEYWORDS EST.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 91)
Hiller, L., Lennon, G., Becker, M., Donald, M.F., Chippelli, B.,
Chissoe, S., Dietrich, N., Dubuque, T., Favell, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohling, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

TITLE
JOURNAL
MEDLINE 97044478
COMMENT On May 18, 1995 this sequence version replaced gi:811133.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40M13 fwd. from Amersham.

FEATURES
source
1..91
/organism="Homo sapiens"
/db_xref="GDB:3813427"
/db_xref="taxon:9606"
/clone="IMAGE:510094"
/clone_lib="StrataGene colon (#937204)"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: colon; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. T-84 colonic epithelial cell line. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor
sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5'
CTCGAGTTTATTTTATTTTATTTT 3' "

BASE COUNT 16 a 35 c 20 g 20 t
ORIGIN

Query Match 0.1%; Score 44; DB 27; Length 91;

Best Local Similarity 100.0%; Pred. No. 0.0029;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29813 GAGGTGAGGTTCAGTGCAGCCGAGATGGCGCCACTGACTCCA 29856
|||||
Db 47 GAGGTGAGGTTCAGTGCAGCCGAGATGGCGCCACTGACTCCA 4

RESULT 4
T25653/3 93 bp mRNA EST 31-OCT-1994
LOCUS
DEFINITION EST00522 Equalized CDNA library from Ltk- cultured cells Mus
musculus CDNA clone A0001D11 3', mRNA sequence.

ACCESSION T25653 GI:562974
VERSION
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS 1 (bases 1 to 93)
TITLE An 'equalized CDNA library' by the reassociation of short
double-stranded cDNAs
JOURNAL Nucleic Acids Res. 18, 5705-5711 (1990)
MEDLINE
COMMENT Contact: KO MSH
Center for Molecular Medicine and Genetics
Wayne State University
5047 Guillen Mall, Detroit, MI 48202
Tel: 3135776708
Fax: 3135776200
Email: msk@cmh.biosci.wayne.edu
Seq primer: T7.

FEATURES
source
1..93
/organism="Mus musculus"
/strain="C3H/An"
/db_xref="taxon:10090"
/clone="A0001D11"
/clone_lib="Equalized CDNA library from Ltk- cultured
cells"

/note="Vector: pBluescript SK(-); Site_1: NotI; Site_2:
EcoRI; mRNA was purified from the mouse Ltk- fibroblastoid
cultured cells. Double-stranded cDNAs were synthesized
from the mRNA using an oligo(dT)-NotI primer. After
shearing to 200-400 bp, a synthetic linker-primer, which
has one blunt and one sticky end and an internal EcoRI
site, was ligated to the cDNAs. The cDNAs were amplified
by the polymerase chain reaction (PCR) using the ligated
linker-primer sequence. After denaturation and
reassociation of the ds-cDNAs, and isolation of single-
stranded cDNAs by hydroxyapatite chromatography, the
single-stranded cDNAs were again amplified by PCR. The
cDNAs were digested with EcoRI and NotI, and inserted
into a plasmid vector."

BASE COUNT 61 a 6 c 4 g 22 t
ORIGIN

Query Match 0.1%; Score 43; DB 20; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.0057;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21515 ATTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 21557
|||||
Db 70 ATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 28

RESULT 5
R18531/3 96 bp mRNA EST 14-APR-1995
LOCUS
DEFINITION yf96e06.r1 Soares Infant brain INTB Homo sapiens CDNA clone

REFERENCE	AUTHORS
TITLE	JOURNAL MEDLINE
COMMENT	
FEATURES	SOURCE
BASE COUNT	ORIGIN
Query Match	Best Local Similarity
Matches	Conservative
QY	D8
LOCUS	AA669571
DEFINITION	ACCESSION VERSION KEYWORDS SOURCE ORGANISM
REFERENCE	AUTHORS
TITLE	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 104)

Hallier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chapellin,B., Chissore,S., Dietrich,N., Dubugue,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Marquis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevisan,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Maria.M.

Generation and analysis of 280,000 human expressed sequence tags Genome Res. 6 (9), 807-828 (1996)

On Sep 12, 1996 this sequence version replaced gi:1404721.

Contact: Wilson Rk Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: -28M13 rev2 from Arersham High quality sequence set: 92.

Location/Qualifiers

1..104

/organism="Homo sapiens"

/db_xref="GDB:4594175"

/db_xref="taxon:9606"

/clone="IMAGE:565111"

/clone_lib="Stratagene lung carcinoma 937218"

/tissue_type="lung carcinoma"

/cell_line="NCI-H69"

/dev_stage="cell line NCI-H69"

/lab_host="SOLR (kanamycin resistant)"

/note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Small cell carcinoma cell line NCI-H69. Average insert size: 1.0 kb; Uni-ZAP XR vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTTTTT 3'"

15 a 28 c 26 g 35 t

Query Match 0.1%; Score 41; DB 29; Length 104;

Best Local Similarity 100.0%; Pred. No. 0.021;

Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 25676 TGATCTGGCTCACGTGCACCTCCCGCTTGCTGGTCAAG 25716
|||||
|||

D8 50 TGATCTGGCTCACGTGCACCTCCCGCTTGCTGGTCAAG 90

RESULT 7

LOCUS AA669571 95 bp mRNA EST 20-NOV-1997

DEFINITION ac18d10.s1 Stratagene ovary (#937217) Homo sapiens cDNA clone IMAGE:856819 3' similar to contains Alu repetitive element;; mRNA sequence.

ACCESSION AA669571

VERSION AA669571.1 GI:2631070

KEYWORDS EST.

SOURCE human.
Human sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 95)

Hallier,L., Allen,M., Bowles,L., Dubugue,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theisinger,B., White,Y., Wylie,T., Waterston,R. and Wilson,R. WashU-NCI Human EST Project

JOURNAL Unpublished (1997)
COMMENT On Nov 6, 1997 this sequence version replaced g1:932677.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amersham.

FEATURES

source

Location/Qualifiers

1..95
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:856819"
/clone_lib="Stratagene ovary (#937217)"
/sex="female"
/dev_stage="49 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: ovary; Vector: Bluescript SK; Site_1: EcoRI;
Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt.
Total ovary tissue, normal, caucasian. Average insert
size: 0.8 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5'
GAATTCGCGACGAG 3' -3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTTCTTTT 3'."

BASE COUNT

15 a 30 c 24 g 26 t

ORIGIN

Query Match 0.1%; Score 40; DB 37; Length 95;
Best Local Similarity 100.0%; Pred. No. 0.045; Indels 0; Gaps 0;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25676 TGATCTGGCTCACTGCACCTCCGCTCTGCGTTCAAG 25715
|||||
Db 30 TGATCTGGCTCACTGCACCTCCGCTCTGCGTTCAAG 69

RESULT 8

T63772

LOCUS T63772 105 bp mRNA EST 17-FEB-1995
DEFINITION YC21F01.s1 Stratagene lung (#937210) Homo sapiens cDNA clone
IMAGE:81337 3' similar to contains Alu repetitive element; mRNA
sequence.
T63772.1 GI:667637

ACCESSION

T63772.1 GI:667637

VERSION

EST.

KEYWORDS

human.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 105)

Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,
Chisoso, S., Dietrich, N., Dubaque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Scheilender, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevisan, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.
and Marra, M.

TITLE

Generation and analysis of 280,000 human expressed sequence tags

JOURNAL

Genome Res. 6 (9), 807-828 (1996)

MEDLINE

9704478

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
Insert Size: 513
High quality sequence stops: 69 source: IMAGE Consortium, LNL This
clone is available royalty-free through LNL; contact the IMAGE
Consortium (info@image.lnl.gov) for further information.

Insert Length: 513 Std Error: 0.00
Seq primer: -21m13
High quality sequence stop: 69.
Location/Qualifiers
1..105

FEATURES

source

/organism="Homo sapiens"
/db_xref="GDB:484954"
/db_xref="taxon:9606"
/clone_image="IMAGE:8137"
/clone_lib="Stratagene lung (#937210)"
/sex="male"
/dev_stage="72 years"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. normal lung. Average insert size: 1.0 kb;
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGCGACGAG
3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3'."

BASE COUNT

17 a 34 c 29 g 25 t

ORIGIN

Query Match 0.1%; Score 40; DB 21; Length 105;
Best Local Similarity 100.0%; Pred. No. 0.043; Indels 0; Gaps 0;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3372 ACCTCGCTCTCCGCGTTCAAGGATTCCTGCTCAGC 3411
|||||
Db 36 ACCTCGCTCTCCGCGTTCAAGGATTCCTGCTCAGC 75

RESULT 9

AA916300

LOCUS AA916300 92 bp mRNA EST 17-JUN-1998
DEFINITION on22c08.s1 NCI-CGAP Lu5 Homo sapiens cDNA clone IMAGE:1557422 3'
similar to contains Alu repetitive element; mRNA sequence.
AA916300
AA916300.1 GI:3055692

ACCESSION

AA916300.1 GI:3055692

VERSION

EST.

KEYWORDS

human.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 92)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On May 5, 1995 this sequence version replaced g1:797742.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

TITLE

CDNA Library Preparation: M. Bento Soares, Ph.D.

AUTHORS

DNA Sequencing by: Greg Lennon, Ph.D.

JOURNAL

Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/dbp/image/image.html

COMMENT

Insert Length: 1519 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 61.
Location/Qualifiers
1..92

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:1557422"
/clone_lib="NCI-CGAP Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a

modified polylinker; 1st strand cDNA was prepared from neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 15 a 34 c 21 g 22 t
 ORIGIN
 Query Match 0.1%; Score 39; DB 40; Length 92;
 Best Local Similarity 100.0%; Pred. No. 0.092;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18014 CACCTCCGGGTTCAGCAATCTCTGCTCAGCCTCC 18052
 DB 45 CACCTCCGGGTTCAGCAATCTCTGCTCAGCCTCC 83

RESULT 10
 LOCUS AA583697 101 bp mRNA EST 26-SEP-1997
 DEFINITION nm38f10.s1 NCI-CGAP_K16 Homo sapiens cDNA IMAGE:1088107 3'
 Similar to contains Alu repetitive element; mRNA sequence.
 AA583697
 ACCESSION AA583697.1 GI:268306
 VERSION EST.
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 101)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Apr 14, 1993 this sequence version replaced gi:692709.
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Stratagene, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 1890 Std Error: 0.00
 Seg primer: -40m13 fwd. ET from Amerham
 High quality sequence stop: 93.
 FEATURES
 Location/Qualifiers
 1..101
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1088107"
 /clone_id="NCI-CGAP_K16"
 /sex="mixed"
 /tissue_type="kidney tumor"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: Kidney; Vector: Bluescript SK-; Site: 1;
 Ecorr: Site: 2; XhoI; Cloned unidirectionally. Primer:
 Oligo dt. Pooled kidney tumors. 5' adaptor sequence: 5'
 GAATTCGGCAGCAG 3' 3' adaptor sequence: 5'
 CTCGATGTTTTTTTTTTTTT 3' Average insert size: 1.0 kb."

BASE COUNT 21 a 29 c 28 g 23 t
 ORIGIN

Query Match 0.1%; Score 39; DB 35; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.087;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 19733 TCAGCGCTAATCCAGCACTTGGGAGGCAAGGCGG 19771
 DB 101 TCAGCGCTAATCCAGCACTTGGGAGGCAAGGCGG 63

RESULT 11
 LOCUS N98204 94 bp mRNA EST 18-NOV-1996
 DEFINITION 0288C3 czappd2.1, Debopam Chakrabarti Plasmodium falciparum cDNA
 clone PF0288C, mRNA sequence.
 ACCESSION N98204
 VERSION N98204.1 GI:1674570
 KEYWORDS EST.
 SOURCE malaria parasite P. falciparum.
 ORGANISM Plasmodium falciparum
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 REFERENCE 1 (bases 1 to 94)
 Dame, J.B., Arnott, D.E., Bourke, P., Chakrabarti, D., Christodoulou, Z.,
 Coppel, R., Cowman, A., Craig, A., Fischer, K., Foster, J., Goodman, N.,
 Hilderberg, K., Holder, A.A., Holt, D., Kemp, D., Lanzer, M., Lim, A.,
 Newbold, C., Ravetch, J.V., Reddy, G.R., Rubio, J., Schuster, S.M.,
 Su, X.-Z., Thompson, J.K., Vital, F., Wellens, T.E. and Werner, E.
 Current status of the Plasmodium falciparum genome project
 Mol. Biochem. Parasitol. 79, 1-12 (1996)
 97001675
 JOURNAL
 MEDLINE
 COMMENT On May 5, 1995 this sequence version replaced gi:797974.
 Contact: Debopam Chakrabarti
 Department of Molecular Biology and Microbiology
 University of Central Florida
 Orlando, FL 32816-2360
 Tel: 407 384 2061
 Fax: 407 384 3095
 Email: dchakr@pegasus.cc.ucf.edu
 Seq primer: 73.
 FEATURES
 Location/Qualifiers
 1..94
 /organism="Plasmodium falciparum"
 /strain="Dd2"
 /db_xref="taxon:5833"
 /clone="PF0288C"
 /clone_id="czappd2.1, Debopam Chakrabarti"
 /lab_host="E. coli XL-1 blue"
 /note="Vector: Lambda ZAP II; Site: 1; Ecorr: I; Site: 2; Xho
 I; POLA+ RNA. from asynchronous blood stage parasites of
 the Dd2 isolate cultured in vitro. was reverse transcribed
 using an oligo dt-Xho I primer. Second strand was
 prepared using RNase H and DNA polymerase I. Ecorr I
 adaptors were ligated to the cDNA, and it was digested
 with Xho I. Prepared fragments were ligated into Ecorr I +
 Xho I digested lambda ZAP II vector."

BASE COUNT 49 a 4 c 6 g 31 t 4 others
 ORIGIN

Query Match 0.1%; Score 38; DB 25; Length 94;
 Best Local Similarity 100.0%; Pred. No. 0.18;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21527 ATTATATTATTATTATTATTATTATTATTATT 21564
 DB 64 ATTATATTATTATTATTATTATTATTATTATT 27

RESULT 12
 LOCUS C05051 104 bp mRNA EST 30-JUL-1996
 DEFINITION C05051 Human heart cDNA (Ynakamura) Homo sapiens cDNA clone
 3NC4544, mRNA sequence.
 ACCESSION C05051
 VERSION C05051.1 GI:1468302
 KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 104)
Tanaka,T., Ogiwara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and Nakamura,Y.
TITLE Construction of a normalized directionally cloned cDNA library from adult heart and analysis of 3040 clones by partial sequencing
JOURNAL Genomics 35 (1), 231-235 (1996)
MEDLINE 96299762
COMMENT Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.
location/Qualifiers
1..104
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="3NHc4544"
/clone_1lb="Human heart cDNA (Ynakamura)"
/dev_stage="adult"
/note="Organ: heart; normalized directionally cloned cDNA from adult heart"

BASE COUNT 24 a 26 c 37 g 17 t
ORIGIN

Query Match 0.1%; Score 38; DB 27; Length 104;
Best Local Similarity 100.0%; Pred. No. 0.17; Mismatches 0; Gaps 0;
Matches 38; Conservative 0; Indels 0;

Qy 3378 GCGTCCCGGTTCAAGCATCTCTCGCTCAGCCTCC 3415
|||||
Db 38 GCGTCCCGGTTCAAGCATCTCTCGCTCAGCCTCC 1

RESULT 13
AA548135 92 bp mRNA EST 14-AUG-1997
LOCUS nK55a11.s1 NCI_CGAP_P7 Homo sapiens cDNA clone IMAGE:1017404, MRYA
DEFINITION
ACCESSION AA548135
VERSION AA548135.1 GI:2318417
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 92)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1397885.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: David G. Bostwick, M.D., Rodrigo F. Chuang, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bbrp/image/image.html

JOURNAL
COMMENT
On Sep 12, 1996 this sequence version replaced gi:1397885.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: David G. Bostwick, M.D., Rodrigo F. Chuang, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
www-bio.lnl.gov/bbrp/image/image.html

Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 70.
Location/Qualifiers

FEATURES

source 1..92
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1017404"
/clone_1lb="NCI_CGAP_P7"
/sex="male"
/tissue_type="low-grade prostatic neoplasia"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pAMP10; mRNA made from prostate intraepithelial neoplasia (high-grade), cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp."

BASE COUNT 30 a 18 c 27 g 17 t
ORIGIN

Query Match 0.1%; Score 37; DB 35; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.37; Mismatches 0; Gaps 0;
Matches 37; Conservative 0; Indels 0;

Qy 18257 TTTTGTGAGACAGAGCTCAGCTGTCACCCAGCT 18293
|||||
Db 86 TTTTGTGAGACAGAGCTCAGCTGTCACCCAGCT 50

RESULT 14
B80126/c 102 bp DNA GSS 24-OCT-1998
LOCUS CIT-HSP-2042J24.TR CIT-HSP Homo sapiens genomic clone 2042J24, genomic survey sequence.
ACCESSION B80126
VERSION B80126.1 GI:2867149
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 102)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)
Unpublished (1998)
Other_GSS: CIT-HSP-2042J24.TF
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tifg.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tifg.org/tidb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.
location/Qualifiers
1..102
/organism="Homo sapiens"
/db_xref="GDB:705355"
/db_xref="taxon:9606"
/clone="2042J24"
/clone_1lb="CIT-HSP"
/sex="male"
/cell_type="Sperm"
/note="Vector: pBelOBAC1; Site_1: HindIII; Site_2: HindIII"

BASE COUNT 20 a 32 c 19 g 31 t
ORIGIN

Query Match 0.1%; Score 37; DB 81; Length 102;
 Best Local Similarity 100.0%; Pred. No. 0.35;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: March 31, 2000, 13:42:06
 Job time: 210380 sec

QY 21782 AGCCTCCCAAGTGTGGATTACAGCGGTAGCCAC 21818
 |||
 Db 102 AGCCTCCCAAGTGTGGATTACAGCGGTAGCCAC 66

RESULT 15
 T94466
 LOCUS

DEFINITION T94466 105 bp mRNA EST 24-MAR-1995
 Y635B02.r1 StrataGene lung (#937210) Homo sapiens cDNA clone
 IMAGE:119691 5' similar to contains Alu repetitive element; mRNA
 sequence.

ACCESSION T94466
 VERSION T94466.1 GI:727954
 KEYWORDS EST.

SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 105)
 AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,
 Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
 Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,
 Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
 Rohlfing, T., Schellenger, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
 Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
 and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (9), 807-828 (1996)
 MEDLINE 97044478

COMMENT Contact: Wilson R
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

Insert Size: 614
 High quality sequence stops: 70 Source: IMAGE Consortium, LIND. This
 clone is available royalty-free through LIND; contact the IMAGE
 Consortium (info@image.llnl.gov) for further information.
 Insert length: 614 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 70.
 Location/Qualifiers

FEATURES
 source 1..105
 /organism="Homo sapiens"
 /db_xref="GDB:487980"
 /db_xref="taxon:9606"
 /clone="IMAGE:119691"
 /clone_lib="Stratagene lung (#937210)"
 /sex="male"
 /dev_stage="72 years"
 /lab_host="SOLR cells (Kanamycin resistant)"
 /note="Organ: lung; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
 Oligo dt. normal lung. Average insert size: 1.0 kb;
 Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGACGAG
 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'."

BASE COUNT 31 a 20 c 34 g 18 t 2 others
 ORIGIN

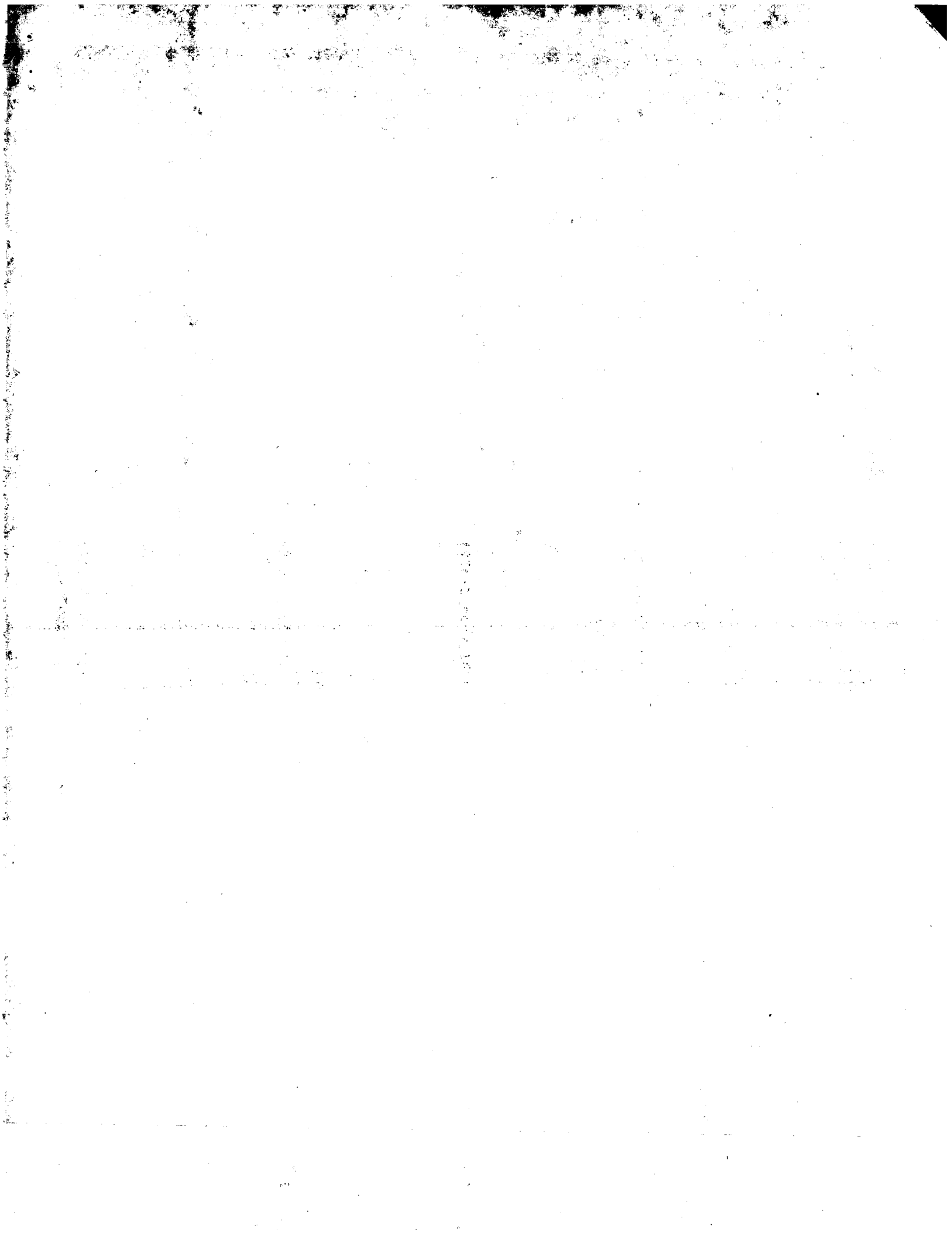
Query Match 0.1%; Score 37; DB 21; Length 105;
 Best Local Similarity 100.0%; Pred. No. 0.34;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19898 CTGAGCGAGAGATCTTGAACCGAGAGCGAGAG 19934
 |||
 Db 57 CTGAGCGAGAGATCTTGAACCGAGAGCGAGAG 93

Mon Apr 3 08:24:12 2000

us-08-852-495c-1_copy_165000_195000.rst

Page 9



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2000, 17:25:40 ; Search time 13412.4 Seconds

(without alignments) 1 updates/sec
-6791.801 Million cell updates/sec

Title: US-08-852-495C-1_COPY_190000_220000

Perfect score: 30001
Sequence: 1 GATATATCTCATGTGTAATG.....AGGGCGCAAGAGATCATG 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 82193 seqs, -1518192014 residues

Word size : 0

Total number of hits satisfying chosen parameters: 345848

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_ov:*
5: gb_ov:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pl4:*
11: gb_pl5:*
12: gb_pl6:*
13: gb_pl7:*
14: gb_pl8:*
15: gb_pl9:*
16: gb_pl10:*
17: gb_pl11:*
18: gb_pl12:*
19: gb_pl13:*
20: gb_pl14:*
21: gb_pl15:*
22: gb_pl16:*
23: gb_pl17:*
24: gb_pl18:*
25: gb_pl19:*
26: gb_pl20:*
27: gb_pl21:*
28: gb_pl22:*
29: gb_pl23:*
30: gb_pl24:*
31: gb_pl25:*
32: gb_pl26:*
33: gb_pl27:*
34: gb_pl28:*
35: gb_pl29:*
36: gb_pl30:*
37: gb_pl31:*
38: gb_pl32:*
39: gb_pl33:*
40: gb_pl34:*
41: gb_pl35:*
42: gb_pl36:*
43: gb_pl37:*

44: gb_htg6:*
45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hum5:*
50: gb_pl3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	0.2	60	5 I31332	I31332 Sequence 24
2	50	0.2	91	5 I31254	I31254 Sequence 16
3	50	0.2	92	5 I31518	I31518 Sequence 43
4	46	0.2	66	40 AF087511	AF087511 Homo sapi
5	43	0.1	65	9 HUMALVNCB	L36836 Homo sapien
6	42	0.1	56	5 I31153	I31153 Sequence 65
7	41	0.1	50	5 I31467	I31467 Sequence 37
8	40	0.1	40	5 A68621	A68621 Sequence 1
9	40	0.1	85	5 A45374	A45374 Sequence 44
10	40	0.1	85	5 AR061179	AR061179 Sequence
11	40	0.1	85	5 HSWG3B103	X86915 H. sapiens s
12	37	0.1	60	5 AR051487	AR051487 Sequence
13	37	0.1	76	5 AR051499	AR051499 Sequence
14	36	0.1	80	9 HUMBRKFAC	M36133 Human alpha
15	36	0.1	101	10 S79560	S79560 HRX (intron
16	36	0.1	101	10 S79561	S79561 dHRX (partl
17	35	0.1	33	5 A25212	A25212 Inter-Alu s
18	35	0.1	35	24 E09140	E09140 Synthetic D
19	35	0.1	60	11 HSCBFB5108	AF084962 Homo sapi
20	35	0.1	68	13 HUMUT997A	L39153 Human STS U
21	35	0.1	83	13 G37817	G37817 SAH Plasmid
22	35	0.1	85	5 AR051522	AR051522 Sequence
23	35	0.1	90	9 HUMLDLRF	K03555 Human low d
24	34	0.1	41	9 HUMALVNCB	L36835 Homo sapien
25	34	0.1	67	9 HUMALVNCB	L36843 Homo sapien
26	34	0.1	72	13 G31952	G31952 MCM92 Chick
27	34	0.1	75	3 BRIG445	X85061 B. taurus co
28	34	0.1	88	3 OADU5702	U15702 Ovis aries
29	34	0.1	95	13 AU026531	AU026531 Rattus no
30	34	0.1	95	13 AU026788	AU026788 Rattus no
31	34	0.1	99	13 G37948	G37948 C4W78 Plasm
32	34	0.1	99	13 G37948	G37948 C4W78 Plasm
33	34	0.1	101	13 G37906	G37906 C3M4 Plasm
34	33	0.1	33	5 I31400	I31400 Sequence 31
35	33	0.1	40	5 I31263	I31263 Sequence 17
36	33	0.1	46	5 I31159	I31159 Sequence 71
37	33	0.1	51	5 I31405	I31405 Sequence 31
38	33	0.1	54	5 AR002288	AR002288 Sequence
39	33	0.1	54	5 AR053139	AR053139 Sequence
40	33	0.1	60	34 DBOR81	M36564 D. melanogas
41	33	0.1	62	3 I07633	I07633 Sequence 2
42	33	0.1	79	10 S73203	S73203 AL1-1 (tand
43	33	0.1	82	3 HUMRSCB	K02284 Human poly
44	33	0.1	83	13 G37817	G37817 SAH Plasmid
45	33	0.1	87	9 HUML	L23536 Human Chrom

ALIGNMENTS

RESULT 1

I31332 I31332 60 bp DNA

LOCUS I31332 Sequence 244 from patent US 5582979.

DEFINITION I31332

ACCESSION I31332

VERSION I31332.1 GI:1822123

PAT 06-FEB-1997

```

KEYWORDS      Unknown.
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE      1 (bases 1 to 60)
AUTHORS        Weber,J.L.
TITLE          Length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and
               method of using the same
JOURNAL        Patent: US 5582979-A 244 10-DEC-1996;
FEATURES       Location/Qualifiers
               source
                   1..60
                       /organism="unknown"
BASE COUNT     18 a      30 c      0 g      12 t
ORIGIN
Query Match    0.2%; Score 54; DB 5; Length 60;
Best Local Similarity 100.0%; Pred.No. 1.3e-15;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3932 CTTCTCTCTCTCTCTCTCTCTCACACACACACACACACACACACACA 3985
Db 2 CTTCTCTCTCTCTCTCTCTCTCACACACACACACACACACACACACACA 55
|||||
RESULT 2
I31254 LOCUS I31254 91 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 166 from patent US 5582979.
ACCESSION I31254
VERSION I31254.1 GI:1822045
KEYWORDS
SOURCE .
ORGANISM Unknown.
UNCLASSIFIED.
REFERENCE 1 (bases 1 to 91)
AUTHORS Weber,J.L.
TITLE Length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and
method of using the same
JOURNAL Patent: US 5582979-A 166 10-DEC-1996;
FEATURES Location/Qualifiers
source 1..91
            /organism="unknown"
BASE COUNT 20 a      43 c      1 g      27 t
ORIGIN
Query Match    0.2%; Score 50; DB 5; Length 91;
Best Local Similarity 100.0%; Pred.No. 1.3e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3932 CTTCTCTCTCTCTCTCTCTCTCACACACACACACACACACACACA 3981
Db 30 CTTCTCTCTCTCTCTCTCTCTCACACACACACACACACACACACACA 79
|||||
RESULT 3
I31518 LOCUS I31518 92 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 430 from patent US 5582979.
ACCESSION I31518
VERSION I31518.1 GI:1822309
KEYWORDS
SOURCE .
ORGANISM Unknown.
UNCLASSIFIED.
REFERENCE 1 (bases 1 to 92)
AUTHORS Weber,J.L.
TITLE Length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and
method of using the same
JOURNAL Patent: US 5582979-A 430 10-DEC-1996;
FEATURES Location/Qualifiers
source 1..92

```

[illegible]

AUTHORS Paabo,S.
TITLE Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 86 (6), 1939-1943 (1989)
MEDLINE 89184542
FEATURES

source location/Qualifiers
1. .65
/organism="Homo sapiens"
/isolate="4000 year old remains from Neht-ankh"
/db_xref="taxon:9606"
/tissue_type="liver"
repeat_region <1. .>65
/rpt_family="Alu"
BASE COUNT 17 a 17 c 21 g 10 t
ORIGIN

Query Match 0.1%; Score 43; DB 9; Length 65;
Best Local Similarity 100.0%; Pred. No. 4.2e-10;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24735 GACCTCGATCCGCCGCTCGCCCAAGTCTGGAT 24777
Db 65 GACCTCGATCCGCCGCTCGCCCAAGTCTGGAT 23

RESULT 6
LOCUS 131153 56 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 65 from patent US 5582979.
ACCESSION 131153
VERSION 131153.1 GI:1821944
KEYWORDS
SOURCE Unknown.
ORGANISM Unidentified.

REFERENCE 1 (bases 1 to 56)
AUTHORS Weber,J.L.

TITLE Length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and
JOURNAL method of using the same
PATENT: US 5582979-A 65 10-DEC-1996;
FEATURES location/Qualifiers
source 1. .56
/organism="unknown"

BASE COUNT 23 a 28 c 0 g 5 t
ORIGIN

Query Match 0.1%; Score 42; DB 5; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3944 CTCCTCTCTCACACACACACACACACACACA 3985
Db 1 CTCCTCTCTCACACACACACACACACACACACA 42

RESULT 7
LOCUS 131467 50 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 379 from patent US 5582979.
ACCESSION 131467
VERSION 131467.1 GI:1822258
KEYWORDS
SOURCE Unknown.
ORGANISM Unidentified.

REFERENCE 1 (bases 1 to 50)
AUTHORS Weber,J.L.
TITLE Length polymorphisms in (dc-da).sub.n.(dg-dt).sub.n sequences and
JOURNAL method of using the same
PATENT: US 5582979-A 379 10-DEC-1996;
FEATURES location/Qualifiers

source 1. .50
/organism="unknown"
BASE COUNT 20 a 25 c 0 g 5 t
ORIGIN

Query Match 0.1%; Score 41; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.3e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3945 TCTCTCTCACACACACACACACACACACACA 3985
Db 1 TCTCTCTCACACACACACACACACACACACA 41

RESULT 8
LOCUS A68621/c A68621 40 bp DNA PAT 06-MAY-1999
DEFINITION Sequence 1 from Patent WO9801573.
ACCESSION A68621
VERSION A68621.1 GI:4759648
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.

REFERENCE 1 (bases 1 to 40)
AUTHORS Resnick M.A., Laktionov V.L., Koupina,N.Y. and Perkins,E.L.
TITLE TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
JOURNAL Patent: WO 9801573-A 15-JAN-1998;
US HEALTH (US) location/Qualifiers

FEATURES source 1. .40
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 7 a 12 c 13 g 8 t
ORIGIN

Query Match 0.1%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24755 TCGGCTCCCAAGTCTGGATTACAGCGTGAACCC 24794
Db 40 TCGGCTCCCAAGTCTGGATTACAGCGTGAACCC 1

RESULT 9
LOCUS A45374/c A45374 85 bp DNA PAT 07-MAR-1997
DEFINITION Sequence 44 from Patent WO9517522.
ACCESSION A45374
VERSION A45374.1 GI:2299846
KEYWORDS
SOURCE unidentified.

ORGANISM unidentified.
REFERENCE 1 (bases 1 to 85)
AUTHORS Jeffreys,A.J. and Armour,J.
TITLE IDENTIFICATION OF SIMPLE TANDEM REPEATS
JOURNAL Patent: WO 9517522-A 44 29-JUN-1995;
UNIV LEICESTER (GB) location/Qualifiers
COMMENT Other publication AU 1277995 950710.
FEATURES source 1. .85
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 58 a 1 c 24 g 2 t
ORIGIN

Query Match 0.1%; Score 40; DB 5; Length 85;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;

KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 80)
Kornreich, R., Bishop, D.F. and Desnick, R.J.
Alpha-galactosidase A gene rearrangements causing Fabry disease:
identification of short direct repeats at breakpoints in an
Alu-rich gene

JOURNAL J. Biol. Chem. 265, 9319-9326 (1990)

MEDLINE 90264427

FEATURES
source location/Qualifiers
1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 18 a 28 c 17 g 17 t
ORIGIN

Query Match 0.1%; Score 36; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18946 GCCTCGGCTCCCAAGTGTGAGATTACAGCATG 18981

Db 16 GCCTCGGCTCCCAAGTGTGAGATTACAGCATG 51

RESULT 15
S79560 579560 101 bp DNA PRI 27-JAN-1996
LOCUS HRX [intron 6] [human, genomic, 101 nt].
DEFINITION S79560
ACCESSION S79560
VERSION S79560.1 GI:1168041
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 101)
Bernard, O.A., Romana, S.P., Schichman, S.A., Mauchauffe, M.,
Jonveaux, P. and Berger, R.
Partial duplication of HRX in acute leukemia with trisomy 11
leukemia 9 (9), 1487-1490 (1995)
95387660
GenBank staff at the National Library of Medicine created this
entry [NCBI g1bbsq 170933] from the original journal article.
This sequence comes from Fig. 3.
Map location: 11q23.

FEATURES
source location/Qualifiers
1..101
/organism="Homo sapiens"
/db_xref="taxon:9606"

gene 1..101
/partial
/gene="HRX"

BASE COUNT 27 a 21 c 28 g 25 t
ORIGIN

Query Match 0.1%; Score 36; DB 10; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14193 GAGACGAGCTGCGCAACATGTGAACCCGTCTC 14228

Db 11 GAGACGAGCTGCGCAACATGTGAACCCGTCTC 46

KW linkage analysis; genetic disease; animal; plant; breeding; locus;
KM hybridisation; chromosome; ds.
OS Homo sapiens.
PN US5582979-A.
PD 10-DEC-1996.
PF 21-APR-1989; 341562.
PR 21-APR-1989; US-341562.
PR 05-SEP-1991; US-754351.
PR 04-APR-1994; US-222177.
PA (MARS-) MARSHFIELD CLINIC.
PI Weber JL.
DR WPI; 97-042299/04.
PT Detection of polymorphic genetic markers of the form
PT (dc-da)n(dg-dt)n - using novel nucleic acid mols. as primers
PS Disclosure: Column 9-10: 186pp; English.
CC The invention relates to the isolation of polymorphic repeat sequences
CC having the sequence (dc-da)n.(dg-dt)n which can be used as genetic
CC markers. Primers based on these sequences can be used to detect these
CC repeats, especially for use in e.g. paternity or maternity testing,
CC human genetic analysis such as linkage analysis of genetic disease,
CC commercial animal or plant breeding or pedigree analysis.
CC The repeats, when analysed, fall into 4 categories:
CC 1) perfect repeats which are alternating tandem CA repeats with no
CC interruptions and without adjacent repeats of another sequence;
CC 2) imperfect repeats which are defined as 2 or more runs of uninterrupted

Query Match 0.2%; Score 50; DB 1; Length 91;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3332 CTCCTCTCTCTCTCTCTCTCTCACACACACACACACACACACA 3981
DB 30 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACACACACA 79

RESULT 3
T66081
ID T66081 standard; DNA; 92 BP.
AC T66081;
DT 18-JUN-1997 (first entry)
DE (dc-da)n.(dg-dt)n polymorphic repeat sequence #12.
KM Polymorphism; repeat sequence; genetic marker; primer; amplification;
KM PCR; polymerase chain reaction; paternity; maternity; human; pedigree;
KM linkage analysis; genetic disease; animal; plant; breeding; locus;
KM hybridisation; chromosome; ds.
OS Homo sapiens.
PN US5582979-A.
PD 10-DEC-1996.
PF 21-APR-1989; 341562.
PR 21-APR-1989; US-341562.
PR 05-SEP-1991; US-754351.
PR 04-APR-1994; US-222177.
PA (MARS-) MARSHFIELD CLINIC.
PI Weber JL.
DR WPI; 97-042299/04.
PT Detection of polymorphic genetic markers of the form
PT (dc-da)n(dg-dt)n - using novel nucleic acid mols. as primers
PS Example 8: Column 57-58: 186pp; English.
CC The invention relates to the isolation of polymorphic repeat sequences
CC having the sequence (dc-da)n.(dg-dt)n which can be used as genetic
CC markers. Primers based on these sequences can be used to detect these
CC repeats, especially for use in e.g. paternity or maternity testing,
CC human genetic analysis such as linkage analysis of genetic disease,
CC commercial animal or plant breeding or pedigree analysis.
CC The repeats, when analysed, fall into 4 categories:
CC 1) perfect repeats which are alternating tandem CA repeats with no
CC interruptions and without adjacent repeats of another sequence;
CC 2) imperfect repeats which are defined as 2 or more runs of uninterrupted

CC CA repeats separated by no more than 3 consecutive non-repeat bases;
CC 3) compound perfect repeats which are uninterrupted runs of CA separated
CC by no more than 3 consecutive non-repeat bases from a run of at least
CC 5 uninterrupted dinucleotide or longer repeats of a sequence other than
CC (dc-da)n.(dg-dt)n, or from at least 10 uninterrupted mononucleotides; and
CC 4) imperfect compound repeats which are defined as for the perfect
CC compound repeats except that the runs of CA are interrupted.
CC This sequence is an example of a compound imperfect repeat sequence of
CC structure: T(CT)12GTT(TC)11T(CA)14A(AC)6.
SQ Sequence 92 BP; 21 A; 43 C; 1 G; 27 T;

Query Match 0.2%; Score 50; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3332 CTCCTCTCTCTCTCTCTCTCTCACACACACACACACACACACA 3981
DB 30 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACACACACA 79

RESULT 4
T65707
ID T65707 standard; DNA; 56 BP.
AC T65707;
DT 17-JUN-1997 (first entry)
DE Repeat sequence from polymorphic marker clone Mf5.
KM Polymorphism; repeat sequence; genetic marker; primer; amplification;
KM PCR; polymerase chain reaction; paternity; maternity; human; pedigree;
KM linkage analysis; genetic disease; animal; plant; breeding; locus;
KM hybridisation; chromosome; ds.
OS Homo sapiens.
PN US5582979-A.
PD 10-DEC-1996.
PF 21-APR-1989; 341562.
PR 21-APR-1989; US-341562.
PR 05-SEP-1991; US-754351.
PR 04-APR-1994; US-222177.
PA (MARS-) MARSHFIELD CLINIC.
PI Weber JL.
DR WPI; 97-042299/04.
PT Detection of polymorphic genetic markers of the form
PT (dc-da)n(dg-dt)n - using novel nucleic acid mols. as primers
PS Claim 1; Column 9-10: 186pp; English.
CC The invention relates to the isolation of polymorphic repeat sequences
CC having the sequence (dc-da)n.(dg-dt)n which can be used as genetic
CC markers. Primers based on these sequences can be used to detect these
CC repeats, especially for use in e.g. paternity or maternity testing,
CC human genetic analysis such as linkage analysis of genetic disease,
CC commercial animal or plant breeding or pedigree analysis.
CC containing the repeat sequences were isolated by hybridisation of
CC chromosome-specific phage libraries with a synthetic poly(dg-dt)n (dg-dt)
CC probe. Over 100 repeat blocks were isolated. The inserts from the clones
CC were amplified by primers T65798-T66047. Those clones where the repeat
CC sequence has been determined are shown in T65704-797. This repeat
CC sequence is from the marker clone Mf5 which contains the repeat
SQ Sequence 56 BP; 23 A; 28 C; 0 G; 5 T;

Query Match 0.1%; Score 42; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3344 CTCCTCTCTCTCTCTCTCTCTCACACACACACACACACACACA 3985
DB 1 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACACACACA 42

RESULT 5
T65791
ID T65791 standard; DNA; 50 BP.
AC T65791;

DE Repeat sequence from polymorphic marker clone Mfd120.
 KW Polymorphism: repeat sequence; genetic marker; primer: amplification;
 KW PCR; polymerase chain reaction; paternity; maternity; human; pedigree;
 KW linkage analysis; genetic disease; animal; plant; breeding; locus;
 KW hybridisation; chromosome; ds.
 OS Homo sapiens.
 PN US5582979-A.
 PD 10-DEC-1996.
 PR 21-APR-1989; 341562.
 PR 21-APR-1989; US-341562.
 PR 05-SEP-1991; US-754351.
 PR 04-APR-1994; US-222177.
 PA (MARS-) MARSHFIELD CLINIC.
 PI Weber JL;
 DR WPI: 97-042299/04.
 PR Detection of polymorphic genetic markers of the form
 PS (dc-da)n(dg-dt)n - using novel nucleic acid mols. as primers
 PS Claim 1; Column 13-14; 18pp; English.
 CC The invention relates to the isolation of polymorphic repeat sequences
 CC having the sequence (dc-da)n.(dg-dt)n which can be used as genetic
 CC markers. Primers based on these sequences can be used to detect these
 CC repeats, especially for use in e.g. paternity or maternity testing,
 CC human genetic analysis such as linkage analysis of genetic disease,
 CC commercial animal or plant breeding or pedigree analysis. Clones
 CC containing the repeat sequences were isolated by hybridisation of
 CC chromosome-specific phage libraries with a synthetic poly(dc-da).(dg-dt)
 CC probe. Over 100 repeat blocks were isolated. The inserts from the clones
 CC were amplified by primers T65798-T66047. Those clones from the clones
 CC sequence has been determined are shown in T65704-T97. This repeat
 CC sequence is from the marker clone Mfd120 which contains the repeat
 CC sequence having the formula: (TC)5(AC)20.
 SQ Sequence 50 BP; 20 A; 25 C; 0 G; 5 T;

Query Match 0.1%; Score 41; DB 1; Length 50;
 Best Local Similarity 100.0%; Pred. No. 0.00025;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 3945 TCTCTCTCAGACACACACACACACACACACACACACACA 3985
 |||
 Db 1 TCTCTCTCAGACACACACACACACACACACACACACACA 41

RESULT 6
 ID Q33603 standard; DNA; 56 BP.
 AC Q33603;
 DT 02-FEB-1993 (first entry)
 DE Microsatellite sequence from clone AGLA296.
 KM PCR: selection; primers: OPRIIPIM; breeding: cattle; parentage;
 KM genetic mapping; traits: amplification; ss.
 OS Bos taurus.
 PN WO9213102-A.
 PD 06-AUG-1992.
 PD 15-JAN-1992; U00340.
 PR 15-JAN-1994; US-642342.
 PA (GENM-) GENMARK.
 PI Georges M, Massey JM;
 PR Polymorphic bovine DNA markers - used in genetic identification,
 PT gene mapping, and selective breeding
 PS Table 7; Page 170; 517pp; English.
 CC The sequence is that of a bovine microsatellite sequence obtd. by
 CC screening a library of bovine MboI DNA fragments of between
 CC 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.
 CC One out of 50 clones cross-hybridised. Assuming independent
 CC distribution of microsatellites and MboI sites, the frequency of
 CC (T6)n >9 microsatellites in the bovine genome is estimated at >100,
 CC 000. The sequence information for ca. 230 such bovine microsatellites
 CC is summarised in the specification and indexed herein (see below).
 CC The sequences upstream and downstream of the microsatellite sequence
 CC were used to generate the required PCR primers for in vitro

```

CC amplification of the corresp. microsatellite (using the program
CC OPTIPRIM). The microsatellites may be used to identify individuals,
CC for percentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved in the determination of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SQ Sequence 56 Bp; 19 A; 0 C; 28 G; 9 T;

Query Match 0.1%; Score 41; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3932 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACAC 3972
|||||
DB 41 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACAC 1

RESULT 7
Q33779/c
ID Q33779 standard; DNA; 62 Bp.
AC Q33779;
DT 02-FEB-1993 (first entry)
DE Microsatellite sequence from clone TGLA182.
KW PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;
KM genetic mapping; traits; amplification; ss.
OS Bos taurus.
PN WO9213102-A.
PD 06-AUG-1992.
PF 15-JAN-1992; U000340.
PR 15-JAN-1991; US-642342.
PA (GENM-) GENMARK.
PI Georges M. Massey JM;
PR Polymorphic bovine DNA markers - used in genetic identification,
PI gene mapping, and selective breeding
PS Table 7; Page 241; 517pp; English.

CC The sequence is that of a bovine microsatellite sequence obtd. by
CC screening a library of bovine MboI DNA fragments of between
CC 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MboI sites, the frequency of
CC (T6)n >9 microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for in vitro
CC amplification of the corresp. microsatellite (using the program
CC OPTIPRIM). The microsatellites may be used to identify individuals,
CC for percentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved in the determination of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also Q33501-34437.
SQ Sequence 62 Bp; 22 A; 0 C; 31 G; 9 T;

Query Match 0.1%; Score 41; DB 1; Length 62;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3932 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACAC 3972
|||||
DB 41 CTCCTCTCTCTCTCTCTCTCTCTCACACACACACACACAC 1

RESULT 8
Q33594/c
ID Q33594 standard; DNA; 63 Bp.
AC Q33594;
DT 02-FEB-1993 (first entry)
DE Microsatellite sequence from clone AGLA29.
KW PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;
KM genetic mapping; traits; amplification; ss.

```

```
CC      Bos taurus.  
PN      MO92L3102-A.  
PD      06-AUG-1992.  
PF      15-JAN-1992; U00340.  
PR      15-JAN-1991; US-642342.  
PA      (GENM-) GENMARK.  
PI      Georges M. Massey JM;  
PT      MPI: 92-284684/34.  
DR      Polymorphic bovine DNA markers - used in genetic identification,  
PT      gene mapping, and selective breeding  
PS      Table 7; Page 167; 517pp; English.  
CC      The sequence is that of a bovine microsatellite sequence obd. by  
CC      screening a library of bovine Mbcl DNA fragments of between  
CC      250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.  
CC      One out of 50 clones cross-hybridised. Assuming independent  
CC      distribution of microsatellites and Mbcl sites, the frequency of  
CC      (T6)n >9 microsatellites in the bovine genome is estimated at >100,  
CC      000. The sequence information for ca. 230 such bovine microsatellites  
CC      is summarised in the specification and indexed herein (see below).  
CC      The sequences upstream and downstream of the microsatellite sequence  
CC      were used to generate the required PCR primers for In Vitro  
CC      amplification of the corresp. microsatellite (using the program  
CC      OPTIPRM). The microsatellites may be used to identify individuals,  
CC      for parentage testing, and in the genetic mapping of economic trait  
CC      loci, or genes involved the determinism of economically important  
CC      traits esp. in cattle, to allow selective breeding.
```

```
CC See also Q33501-34437.
SQ Sequence 40 BP; 11 A; 0 C; 20 G; 9 T;

Query Match          0.1%; Score 40; DB 1; Length 40.
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3932 CTCCTCTCTCTCTCTCTCTCTCACACACACACACA 3971
      |||||||
DB 40 CTCCTCTCTCTCTCTCTCTCTCACACACACACACA 1

RESULT 10
V19044/C
ID V19044 standard; DNA: 40 BP.
AC V19044;
DT 28-JUL-1998 (first entry)
DE Alu PCR primer 1.
KW PCR; primer; amplification; Alu repeat sequence; vector;
KM circular yeast artificial chromosome; YAC; ss.
OS Synthetic.
SN Saccharomyces sp.
PN WO9801573-A1.
PD 15-JAN-1998.
PF 09-JUL-1996; U11478.
PR 09-JUL-1996; WO-011478.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI Kourkina NY, Lariouov VL, Perkins EL, Resnick MA;
DR WPI: 98-110234/10.
PT Preparation of yeast artificial chromosomes - by in vivo
recombination using vector comprising yeast centromere, marker,
yeast telomere and nucleic acid for recombination
PS Example 1; Page 45; 117pp; English.
CC This is the nucleotide sequence for the PCR primer used in the
CC amplification of the Alu repeat sequence, which is used to
CC demonstrate the processes described in the invention. It involves
CC the creation and use of circular yeast artificial chromosome (YAC)
CC to selectively clone specific nucleic acids from a background of
CC mixed nucleic acids by introducing the vector(s) into E. coli cells.
CC They can be used to rapidly isolate human DNA where only a part of the
CC sequence of DNA is known. Using the methods large fragments of DNA can
CC be easily cloned and analysed.
SQ Sequence 40 BP; 7 A; 12 C; 13 G; 8 T;

Query Match          0.1%; Score 40; DB 1; Length 40.
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24755 TCGGCTCCCCAAGTGTCTGGATTACAGCGGTGACCACC 24794
      |||||||
DB 40 TCGGCTCCCCAAGTGTCTGGATTACAGCGGTGACCACC 1

RESULT 11
Q95218/C
ID Q95218 standard; DNA: 85 BP.
AC Q95218;
DT 08-FEB-1996 (first entry)
DE Simple tandem repeat (STR) corresponding to the 2nd part of wj3b10.
KW Simple tandem repeat; STR; wj3b10; treatment; genetic; diagnosis;
KW characterization; mapping; linkage studies; analysis; alleles;
OS second part; ss.
OS Synthetic.
PN MO9517522-R2.
PD 29-JUN-1995.
PF 21-DEC-1994; G02789.
PR 21-DEC-1993; GB-026052.
PA (UYLE-) UNIV LEICESTER.
PI Armour J, Jeffreys AJ;
DR WPI: 95-240682/31.
PT Identifying simple tandem repeat loci in DNA - by screening DNA
```

[illegible]

```

DE pSPORT-1a synthetic polyester gene 5' fragment insert.
KW Synthetic gene; synthesis; polyester; amino acid analogue; clips;
KW stop codon; stop signal; monomer unit; lactate; glycolate; staples;
KW translation initiation; modified; Met-tRNAf; cyanogen bromide; pins;
KW co-polymer; sutures; drug delivery device; screws; ds.
OS Synthetic.
PN CA2112716-A.
PD 01-JUL-1984.
PF 31-DEC-1993; 112716.
PR 31-DEC-1992; US-999520.
PA (USSU ) US SURGICAL CORP.
PI GruskIn EA.
DR WPI: 94-272324/34.
PT New biocompatible medical devices and articles - using a lactic
PT acid-glycolic acid polyester in which monomer units are
PT individually and specifically determined
PS Example 1, Page 45; 67pp; English.
CC This sequence represents the 5' fragment of the synthetic genes which
CC encode polyesters given in Q71801-02. This fragment was ligated into
CC the Klenow bluntend sequences given in Q71808-09 which represent the 3'
CC and 5' ends of the plasmid pSPORT 1b digested with PstI. This places
CC this synthetic fragment under the control of the E. coli trpA initiator.
CC The E. coli trpA initiator sequence was used to express the sequences
CC given in Q71801-02 in an E. coli cell free medium. The expressed
CC sequences represent synthetic genes which were used in the synthesis
CC of polyesters derived from the alpha-amino acid analogues of an amino acid
CC except Pro. Of the three stop codons, UAA, UAG and UGA, one of these
CC stop codons is reversed for the stop signal for the polyester. The
CC remaining two stop codons are available for encoding the monomer
CC units of the polyester. Lactate and glycolate are encoded by UAA and
CC UAG, respectively. Initiation of translation always starts at an AUG
CC codon, therefore the first lactate or the chain may be encoded by an
CC AUG codon. This is accomplished through the use of a specially modified
CC Met-tRNAf modified to carry lactate instead of Met. Alternatively, a
CC Met is incorporated in the first position of the polymer chain. The
CC resulting polymer is treated with cyanogen bromide to remove the Met.
CC Co-polymers produced in this manner may be used to produce sutures,
CC staples, clips, drug delivery devices, pins and screws.
SO Sequence 80 BP; 47 A; 6 C; 4 G; 23 T;

Query Match          0.1%; Score 39; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches   39; Conservative 0; Mismatches 0; Indels    0; Gaps    0;

QY 10446 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10484
      |||||||
DB 1 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 39

RESULT 14
CJID71814
ID O71814 standard; DNA: 81 BP.
AC O71814;
DT 30-MAR-1995 (first entry)
DE Polyester coding sequence for insertion into pWAL-p2.
KW Synthetic gene; synthesis; polyester; amino acid analogue; clips;
KW stop codon; stop signal; monomer unit; lactate; glycolate; staples;
KW translation initiation; modified; Met-tRNAf; cyanogen bromide; pins;
KW co-polymer; sutures; drug delivery device; screws; ds.
OS Synthetic.
PN CA2112716-A.
PD 01-JUL-1994.
PF 31-DEC-1993; 112716.
PR 31-DEC-1992; US-999520.
PA (USSU ) US SURGICAL CORP.
PI GruskIn EA.
DR WPI: 94-272324/34.
PT New biocompatible medical devices and articles - using a lactic
PT acid-glycolic acid polyester in which monomer units are
PT individually and specifically determined
PS Example 2, Page 53; 67pp; English.
CC The sequences given in Q71814-16 represent fragments of the

```

Query Match	0.1%	Score	39	DB	1	Length	81
Best Local Similarity	100.0%	Pred.	No.	0.00098			
Matches	39	Conservative	0	Mismatches	0	Indels	0
						Gaps	0

RESULT 15

AC Q71809;

synthetic dene: polyester: α -amino acid analogues: c]fns

stop codon: stop signal: monomer unit: lactate: glycolate: staples

KW translation initiation; modified; Met-tRNA^f; cyanogen bromide; plns;

KW co-polymer; sutures; drug delivery device; screws; ds.

OS Synthetic.

FH	Key	Location/Qualifiers
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		

misc_signal

```
ET      d      /*lag=
ET      /note= "trnA initiator"
```

misc feature

FT / *tag= b

ET

```
FT      misc_feature      22.  .93
```

ET

080113716

CA2112/10-A.
01-TMT-100A

FD 002 1334 :
PF 31-DEC-1993 :

PR 31-DEC-1992;

PA (USSU) US SURGICAL CORP.

PI Gruskin EA;

DR WPI: 94-272324/34

PT New biocompatible medical devices and articles - using a lactic

acid glycolic acid polyester in which monomer units are

P1 Individually and specifically detect

This sequence represents a fragment of the plasmid vector pSPORT 1c which is Example 1, page 40, of pp, English.

cc contains the 5' fragment of the synthetic genes which encode polyesters.

CC given in Q71801-02, under the control of the E. coli trpA transcription

CC initiator and followed by FokI and HindIII restriction sites. A further

CC fragment of the synthetic genes were ligated into pSPORT 1c to give

CC plasmid pSPORT Id (see also Q71810-11), with the final fragment being

cc inserted in a third round of ligation to give pSPORI 1e (see also Q/181).

13). The *E. coli* trpA initiator sequence was used to express the full length sequences given in 071801-02 in an *E. coli* cell free medium. The expressed sequences represent synthetic genes which were used in the synthesis of polyesters derived from the α -amino acid analogues of an amino acid except Pro. Of the three stop codons, UAA, UAG and UGA, one of these stop codons is reversed for the stop signal for the polyester. The remaining two stop codons are available for encoding the monomer units of the polyester. Lactate and glycolate are encoded by UAA and UAG, respectively. Initiation of translation always starts at an AUG codon, therefore the first lactate of the chain may be encoded by an AUG codon. This is accomplished through the use of a specially modified Met-tRNA^{fMet} modified to carry lactate instead of Met. Alternatively, a Met is incorporated in the first position of the polymer chain. The resulting polymer is treated with cyanoogen bromide to remove the Met. Co-polymers produced in this manner may be used to produce sutures, staples, clips, drug delivery devices, pins and screws.

```

Query Match 0.1%; Score 39; DB 1; Length 103;
Best Local Similarity 100.0%; Pscore. No. 0.00093;
Matches 39; Conservative 0; Mismatches 0; Gaps 0

```

Qy 10446 AATAAATTAATAATGAATAAAATTAATAATAATAAT 1048
| | | | | | | | | | | | | | | | | |
Db 23 AATAAATTAATAATAATAATAATAATAATAATAAT 61

Search completed: April 1, 2000, 04:18:02
Job time: 260880 sec

Mon Apr 3 08:24:13 2000

us-08-852-495c-1_copy_190000_220000.png

Page 7

Mon Apr 3 08:24:14 2000

us-08-852-495c-1_copy_190000_220000.rml

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2000, 19:15:35 ; Search time 339.54 Seconds
(Without alignments)
10578.471 Million cell updates/sec

Title: US-08-852-495c-1_COPY_190000_220000

Perfect score: 30001

Sequence: 1 GATATATCATGTCGACTG.....AGGGCAAGAGATCATG 30001

Scoring table: OLIGO_NUC

Searched: 214294 seqs, 59861574 residues

Word size : 0

Total number of hits satisfying chosen parameters: 349160

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/lna/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/lna/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/lna/5C_COMB.seq:*
- 4: /cgn2_6/ptodata/1/lna/5D_COMB.seq:*
- 5: /cgn2_6/ptodata/1/lna/6_COMB.seq:*
- 6: /cgn2_6/ptodata/1/lna/PCRU59_COMB.seq:*
- 7: /cgn2_6/ptodata/1/lna/backfilesl.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	0.2	60	1	US-08-222-177A-244 Sequence 244, App
2	50	0.2	91	1	US-08-222-177A-166 Sequence 166, App
3	50	0.2	92	1	US-08-222-177A-430 Sequence 430, App
4	42	0.1	56	1	US-08-222-177A-65 Sequence 65, App
5	41	0.1	50	1	US-08-222-177A-379 Sequence 379, App
6	40	0.1	85	3	US-08-332-766A-44 Sequence 44, App
7	37	0.1	60	3	US-08-454-557C-57 Sequence 57, App
8	37	0.1	60	3	US-08-340-426D-57 Sequence 57, App
9	37	0.1	60	6	US-08-450-673C-57 Sequence 57, App
10	37	0.1	60	6	US-08-450-673C-57 Sequence 57, App
11	37	0.1	76	3	US-08-454-557C-69 Sequence 69, App
12	37	0.1	76	4	US-08-340-426D-69 Sequence 69, App
13	37	0.1	76	4	US-08-450-673C-69 Sequence 69, App
14	37	0.1	76	6	US-08-454-557C-92 Sequence 92, App
15	35	0.1	85	3	US-08-340-426D-92 Sequence 92, App
16	35	0.1	85	4	US-08-450-673C-92 Sequence 92, App
17	35	0.1	85	4	US-08-450-673C-92 Sequence 92, App
18	35	0.1	85	6	US-08-450-673C-92 Sequence 92, App
19	33	0.1	33	1	US-08-222-177A-312 Sequence 312, App
20	33	0.1	40	1	US-08-222-177A-175 Sequence 175, App
21	33	0.1	46	1	US-08-222-177A-71 Sequence 71, App
22	33	0.1	51	1	US-08-222-177A-317 Sequence 317, App
23	33	0.1	54	2	US-08-469-802B-27 Sequence 27, App
24	33	0.1	54	3	US-08-267-803B-45 Sequence 45, App
25	33	0.1	62	7	US-08-222-177A-59 Sequence 59, App
26	32	0.1	33	1	US-08-222-177A-59 Sequence 59, App

27	32	0.1	33	1	US-08-222-177A-218 Sequence 218, App
28	32	0.1	34	1	US-08-222-177A-110 Sequence 110, App
29	32	0.1	34	1	US-08-222-177A-172 Sequence 172, App
30	32	0.1	34	1	US-08-222-177A-180 Sequence 180, App
31	32	0.1	34	1	US-08-222-177A-192 Sequence 192, App
32	32	0.1	34	1	US-08-222-177A-322 Sequence 322, App
33	32	0.1	35	1	US-08-222-177A-77 Sequence 77, App
34	32	0.1	35	1	US-08-222-177A-189 Sequence 189, App
35	32	0.1	36	1	US-08-222-177A-104 Sequence 104, App
36	32	0.1	36	1	US-08-222-177A-358 Sequence 358, App
37	32	0.1	37	1	US-08-222-177A-140 Sequence 140, App
38	32	0.1	38	1	US-08-222-177A-198 Sequence 198, App
39	32	0.1	38	1	US-08-222-177A-397 Sequence 397, App
40	32	0.1	39	1	US-08-222-177A-137 Sequence 137, App
41	32	0.1	40	1	US-08-222-177A-119 Sequence 119, App
42	32	0.1	40	1	US-08-222-177A-400 Sequence 400, App
43	32	0.1	41	1	US-08-222-177A-403 Sequence 403, App
44	32	0.1	41	1	US-08-222-177A-74 Sequence 74, App
45	32	0.1	41	1	US-08-222-177A-183 Sequence 183, App

ALIGNMENTS

RESULT 1
US-08-222-177A-244
; Sequence 244, Application US/08222177A
; Patent No. 5582979
; GENERAL INFORMATION:
; APPLICANT: Weber, James L.
; TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
; TITLE OF INVENTION: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 460
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dewalt Ross & Stevens, S.C.
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,177A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/341,562
; FILING DATE: 21-APR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S.
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: 09865.601
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 831-2106
; TELEFAX: (608) 831-2106
; TELEX:
; INFORMATION FOR SEQ ID NO: 244:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 60 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: mfd67rs
; US-08-222-177A-244
Query Match 0.2%; Score 54; DB 1; Length 60;


```

: APPLICATION NUMBER: US 07/341,562
: FILING DATE: 21-APR-1989
: ATTORNEY/AGENT INFORMATION:
:   NAME: Sara, Charles S.
:   REGISTRATION NUMBER: 30,492
:   REFERENCE/DOCKET NUMBER: 09865, 601
: TELECOMMUNICATION INFORMATION:
:   TELEPHONE: (608) 831-2100
:   TELEFAX: (608) 831-2106
:   TELEX:
: INFORMATION FOR SEQ ID NO: 65:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 56 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: double
:   TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: IMMEDIATE SOURCE:
:   CLONE: mfd5rs
:
: US-08-222-177A-65

```

Query Match	0.1%	Score 42	DB 1	Length 56
Best Local Similarity	100.0%	Pred. No. 4	5e-05	
Matches 42	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	3944	CTCTCTCTCTCACACACACACACACACACACACA	3985	
DB	1	CTCTCTCTCTCACACACACACACACACACACACA	42	

```

1      RESULT      5
2      US-08-222-177A-379
3      ; Sequence 379, Application US/08222177A
4      ; Patent No. 5582979
5      ;
6      ; GENERAL INFORMATION:
7      ;
8      ; APPLICANT: Weber, James L.
9      ; TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
10     ; TITLE OF INVENTION: (dc-da)n.(dc-dt)n SEQUENCES AND METHODS OF USING SAME
11     ; NUMBER OF SEQUENCES: 460
12     ; CORRESPONDENCE ADDRESS:
13     ; ADDRESSEE: Dewitt Ross & Stevens, S.C.
14     ; STREET: 8000 Excelsior Drive, Suite 401
15     ; CITY: Madison
16     ; STATE: Wisconsin
17     ; COUNTRY: USA
18     ; ZIP: 53717-1914
19     ; COMPUTER READABLE FORM:
20     ; MEDIUM TYPE: Floppy disk
21     ; COMPUTER: IBM PC compatible
22     ; OPERATING SYSTEM: PC-DOS/MS-DOS
23     ; SOFTWARE: PatentIn Release #1.0, Version #1.25
24     ; CURRENT APPLICATION DATA:
25     ; APPLICATION NUMBER: US/08/222,177A
26     ; FILING DATE:
27     ; CLASSIFICATION: 435
28     ; PRIOR APPLICATION DATA:
29     ; APPLICATION NUMBER: US 07/341,562
30     ; FILING DATE: 21-APR-1989
31     ; ATTORNEY/AGENT INFORMATION:
32     ; NAME: Sara, Charles S.
33     ; REGISTRATION NUMBER: 30,492
34     ; REFERENCE/DOCKET NUMBER: 09865, 601
35     ; TELECOMMUNICATION INFORMATION:
36     ; TELEPHONE: (608) 831-2100
37     ; TELEFAX: (608) 831-2106
38     ;
39     ; TELEX:
40     ; INFORMATION FOR SEQ ID NO: 379:
41     ; SEQUENCE CHARACTERISTICS:
42     ; LENGTH: 50 base pairs
43     ; TYPE: nucleic acid
44     ; STRANDEDNESS: double
45     ; TOPOLOGY: linear
46     ;

```

```

; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: mfd120rs
US-08-222-177A-379

```

```
Query Match      0.1%; Score 41; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 9.7e-05;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      3945 TCTCTCTCTACACACACACACACACACACACA   3985
         |||||
Db       1 TCTCTCTCTACACACACACACACACACACA   41
```

RESULT 6
US-08-332-766A-44/c
: Sequence 44, Application US/08332766A

; GENERAL INFORMATION:

APPLICANT: JEFFREYS, ALEC J.

10 TITLE OF INVENTION: SIMPLE TANDEM REPEATS

NUMBER OF SEQUENCES: 125

ADDRESSEE: CUSHMAN DABY & CUSHMAN, L.L.P.

STREET: 1100 New

STATE: D. C.

COUNTRY: U.S.A.

COMPUTER READABLE FORM

COMPUTER: IBM PC compatible

```

;
OPERATING SYSTEM:  PC-DOS/MS-DOS

```

```

; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/332,766A

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

FILING DATE: 21-DEC-1993

ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 25,323

REFERENCE/DOCKET NUMBER: 21

TELEPHONE: (202) 861-3000 ;

TELEFAX: (202) 822-0944

; INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

TYPE: nucleic acid

STRANDEDNESS: \$1

MOLECULE TYPE: DNA (genomic)

US-08-332-766A-44

Query match	0.18; score 40; DB 3; Length 85
-------------	---------------------------------

Matches	40;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

[illegible]

RESULT 7
US-08-454-557C-57
; Sequence 57, Application US/08454557C

Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-57

Query Match 0.1%; Score 37; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21188 GCTAATTTTGTATTAGTAGACAGAGGTTTCAC 21224
DB 11 GCTAATTTTGTATTAGTAGACAGAGGTTTCAC 47

RESULT 8
US-08-340-426D-57
Sequence 57, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-57

Query Match 0.1%; Score 37; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21188 GCTAATTTTGTATTAGTAGACAGAGGTTTCAC 21224
DB 11 GCTAATTTTGTATTAGTAGACAGAGGTTTCAC 47

RESULT 9
US-08-450-673C-57
Sequence 57, Application US/08450673C
Patent No. 5948688
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-57

Query Match 0.1%; Score 37; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21188 GCTAATTTTGTATTAGTAGACAGAGGTTTCAC 21224

DB 11 GCTAATTTTGTATTAGAGACAGGTTTCAC 47

RESULT 10

PCT-US95-17111A-57

Sequence 57, Application PC/TUS9517111A

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neural Thread Protein Gene Expression and

TITLE OF INVENTION: Detection of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/17111A

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/340,426

FILING DATE: 14-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609,3840002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 57:

SEQUENCE CHARACTERISTICS:

LENGTH: 60 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

PCT-US95-17111A-57

Query Match 0.1%; Score 37; DB 6; Length 60;

Best Local Similarity 100.0%; Pred. No. 0.0018;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21188 GCTAATTTTGTATTAGAGACAGGTTTCAC 21224

DB 11 GCTAATTTTGTATTAGAGACAGGTTTCAC 47

RESULT 11

US-08-454-557C-69

Sequence 69, Application US/08454557C

Patent No. 5830670

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/454,557C

FILING DATE: 30-MAY-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609,3840003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 69:

SEQUENCE CHARACTERISTICS:

LENGTH: 76 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

US-08-454-557C-69

Query Match 0.1%; Score 37; DB 3; Length 76;

Best Local Similarity 100.0%; Pred. No. 0.0017;

Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21188 GCTAATTTTGTATTAGAGACAGGTTTCAC 21224

DB 11 GCTAATTTTGTATTAGAGACAGGTTTCAC 47

RESULT 12

US-08-340-426D-69

Sequence 69, Application US/08340426D

Patent No. 5948634

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/340,426D

FILING DATE: 14-NOV-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609,3840002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 69:

SEQUENCE CHARACTERISTICS:

LENGTH: 76 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both
US-08-340-426D-69

Query Match 0.1%; Score 37; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21188 GCTAATTTTGTATTTTGTAGACAGAGGTTTCAC 21224
|||||
DB 11 GCTAATTTTGTATTTTGTAGACAGAGGTTTCAC 47

RESULT 13
US-08-450-673C-69
Sequence 69, Application US/08450673C
Patent No. 5948888

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/450,673C

FILING DATE: 30-MAY-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609.3840004

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ. ID NO.: 69:

SEQUENCE CHARACTERISTICS:

LENGTH: 76 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

US-08-450-673C-69

Query Match 0.1%; Score 37; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21188 GCTAATTTTGTATTTTGTAGACAGAGGTTTCAC 21224
|||||
DB 11 GCTAATTTTGTATTTTGTAGACAGAGGTTTCAC 47

RESULT 14
PCT-US95-17111A-69

Sequence 69, Application PC/TUS9517111A

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neural Thread Protein Gene Expression and

TITLE OF INVENTION: Detection of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/17111A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/340,426

FILING DATE: 14-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 0609.3840002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540

INFORMATION FOR SEQ. ID NO.: 69:

SEQUENCE CHARACTERISTICS:

LENGTH: 76 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

PCT-US95-17111A-69

Query Match 0.1%; Score 37; DB 6; Length 76;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21188 GCTAATTTTGTATTTTGTAGACAGAGGTTTCAC 21224
|||||
DB 11 GCTAATTTTGTATTTTGTAGACAGAGGTTTCAC 47

RESULT 15
US-08-454-557C-92

Sequence 92, Application US/08454557C

Patent No. 5850670

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wands, Jack R.

TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/454,557C

FILING DATE: 30-MAY-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203


```

; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 85 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; US-08-454-57C-92

```

```

Query Match          0.1%; Score 35; DB 3; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.0075;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 18907 GCCAGGATGGTCTCGATCTCCTGACCTGCTGATCC 18941
      |||||||||||||||||||||||||||||||||||
Db    9 GCCAGGATGGTCTCGATCTCCTGACCTGCTGATCC 43

```

Search completed: April 1, 2000, 04:07:52
 Job time: 260334 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2000, 13:42:06 ; Search time 8179.17 Seconds
(without alignments)
13849.045 Million cell updates/sec

Title: US-08-852-495c-1_COPY_190000_220000
Perfect score: 30001
Sequence: 1 GTATATATCATGCTACTG.....AGGGCAAAAGATCATG 30001

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4538634 seqs, 1887831982 residues

Word size : 0

Total number of hits satisfying chosen parameters: 130662

Minimum DB seq length: 8

Maximum DB seq length: 105

Post-processing: Listing first 45 summaries

Database : ESR: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *

44: gb_est25: *
45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est31: *
80: gb_est32: *
81: gb_est33: *
82: gb_est34: *
83: em_est31: *
84: em_est32: *
85: em_est33: *
86: em_est34: *
87: gb_est35: *
88: gb_est36: *
89: gb_est37: *
90: gb_est38: *
91: gb_est39: *
92: em_est35: *
93: em_est36: *
94: em_est37: *
95: em_est38: *
96: em_est39: *
97: em_est40: *
98: em_est41: *
99: gb_est40: *
100: gb_est41: *
101: em_est42: *
102: gb_est42: *
103: gb_est43: *
104: gb_est44: *
105: gb_est45: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	48	0.2	102	81	B62983

B62983 CTR-HSP-669

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	B62983/c	B62983 102 bp DNA CIT-HSP-669P15.TP CIT-HSP Homo sapiens genomic clone 669P15.	B62983	B62983.1	GI:2636892	GSS	21-JUN-1998					
2	B48914	B48914 RPII11-4A12 F29274 HSPD19063 H AA151536 VCG7601.S AI305512 qv72D02.X AA053038 z171e12.s AA835205 ak64h01.s AI864984 wK6c11.x AA835411 ak72q06.s AA129957 zn86h04.r AA668571 ac18a10.S R67088 Y130H05.S1 AU072757 AU072757 AI262095 qz28c05.Y AU038446 AU038446 AA0582186 RPII-11-4 B74378 CIT-HSP-202 AI088003 oc24h05.X C25772 C25772 Dict AA614379 np46c03.s AA457423 aa86b02.r D25845 HUMG04217 AI475290 t181c11.x N88930 K668F Huma AA0280224 CITBI-EI- B80126 CIT-HSP-204 AQ319270 RPII11-98 AI973231 w534d06.X AQ276193 CITBI-EI- R70733 Y146G12.r1 AA601314 n015f06.s AA082835 zn21g12.s AA809831 oa40f11.s AA548135 nk55a11.s R09732 YFZ7A12.s1 R33021 YQ11G03.s1 H61099 YR51C12.r1 AA578401 n153C01.s AA494273 ng96b11.s AA228795 nc14e07.s AA807640 nx08b05.s N77004 YV50F05.r1 AA873656 oc02g07.s AI335049 ta44b01.x X88459 H.sapiens D	human.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. Kim, U.-J., Adams, M.D. and Simon, M.I. Determination of clone end sequences of human Bacterial Artificial Chromosomes Unpublished (1997) Contact: Ung-Jin Kim Caltech Genome Research Lab California Institute of Technology Division of Biology, MS 147-75, Pasadena, CA 91125, USA Tel: 626 796 7066 Fax: 626 395 4901 Email: ung@ash.tree.caltech.edu								

ALIGNMENTS

Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/humgen/Bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

Location/Qualifiers
1. 102
/organism="Homo sapiens"
/db_xref="GDB:5491647"
/db_xref="taxon:9606"
/clone_id="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; site_1: HindIII; site_2: HindIII"

BASE COUNT
30 a 30 c 21 g 21 t

ORIGIN

Query Match 0.28; Score 48; DB 81; Length 102;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17313 TTTTGTATTTTGTAGATGGGTTTCACATGTTGCCAGATG 17360
|||||
DB 59 TTTTGTATTTTGTAGATGGGTTTCACATGTTGCCAGATG 12

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
2	B48914/c	B48914 103 bp DNA RPII11-4A12.TP RPII-11 Homo sapiens genomic clone RPII-11-4A12, genomic survey sequence.	B48914	B48914.1	GI:2601151	GSS	08-APR-1999					
3	B48914	B48914 RPII11-4A12 F29274 HSPD19063 H AA151536 VCG7601.S AI305512 qv72D02.X AA053038 z171e12.s AA835205 ak64h01.s AI864984 wK6c11.x AA835411 ak72q06.s AA129957 zn86h04.r AA668571 ac18a10.S R67088 Y130H05.S1 AU072757 AU072757 AI262095 qz28c05.Y AU038446 AU038446 AA0582186 RPII-11-4 B74378 CIT-HSP-202 AI088003 oc24h05.X C25772 C25772 Dict AA614379 np46c03.s AA457423 aa86b02.r D25845 HUMG04217 AI475290 t181c11.x N88930 K668F Huma AA0280224 CITBI-EI- B80126 CIT-HSP-204 AQ319270 RPII11-98 AI973231 w534d06.X AQ276193 CITBI-EI- R70733 Y146G12.r1 AA601314 n015f06.s AA082835 zn21g12.s AA809831 oa40f11.s AA548135 nk55a11.s R09732 YFZ7A12.s1 R33021 YQ11G03.s1 H61099 YR51C12.r1 AA578401 n153C01.s AA494273 ng96b11.s AA228795 nc14e07.s AA807640 nx08b05.s N77004 YV50F05.r1 AA873656 oc02g07.s AI335049 ta44b01.x X88459 H.sapiens D	human.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. Kim, U.-J., Adams, M.D., Rounsley, S.D., Field, C.E., Baas, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C. Use of BAC End Sequences for Sequence-Ready Map Building Unpublished (1997) Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: madams@tigr.org Clones are derived from the human BAC library RPII-11. For BAC library availability, please contact Pieter de Jong (pieter@jlong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/humgen/Bac_end_search/bac_end_search.html Seq primer: SP6 Class: BAC ends.								

Location/Qualifiers
1. 103
/organism="Homo sapiens"
/db_xref="GDB:7501163"
/db_xref="taxon:9606"
/clone="RPII-11-4A12"
/clone_id="RPII-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC63.6; site_1: EcoRI; site_2: EcoRI; RPII11 Human Male BAC Library"

BASE COUNT	30 a	28 c	30 g	15 t
ORIGIN				

Query Match	0.28;	Score 48;	DB 81;	Length 103;
Best Local Similarity	100.00;	Pred. No. 2e-05;		
Matches 48;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 24728 ATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCCTCCCAAAGTGTGGG 24775

Db 53 ATCTCCTGACCTCGTGATCCGCCGCCCTCGGCCCTCCCAAAGTGTGGG 6

RESULT	3	.
F29274/c		
LOCUS	F29274	91 bp mRNA
DEFINITION	HSPD18063 HMJ Homo sapiens cDNA clone s4000074G04, mRNA sequence	EST 13-MAY-1999
ACCESSION	F29274	
VERSION	F29274.1	GI:4814900
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

1 (bases 1 to 91)
Lanfanchi,G., Moraro,T., Caldara,F., Pacchioni,B., Pallavicini,A.
Pandofo,D., Toppo,S., Trevisan,S., Scarso,S. and Valle,G.
Identification of 4370 expressed sequence tags from a
3'-end-specific cDNA library of human skeletal muscle by DNA
sequencing and filter hybridization
Genome Res. 6 (1), 35-42 (1996)
96276048
On Apr 7, 1998 this sequence version replaced g1:3034419.

BASE COUNT	18 a	19 c	38 g	16 t
ORIGIN				

Query Match	0.28;	Score 47;	DB 50;	Length 91;
Best Local Similarity	100.08;	Pred. No. 4.5e-05;		
Matches 47;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 24621 TCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCTGCCACC 24667

Db 63 TCCTGCCTCAGCCTCCCGAGTAGCTGGACTACAGGCGCCTGCCACC 17

RESULT 4
AA415364/C

LOCUS	AAA15364	74 bp	mRNA	EST	16-OCT-1997
DEFINITION	vcg7g01.s1 Knowles Solter mouse 2 cell Mus musculus cdna clone				

ACCESSION	AA415364
VERSION	AA415364.1
	GI:2075370

SOURCE	house mouse.
ORGANISM	Mus musculus.

REFERENCE
1 (bases 1 to 74)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

FEATURES	Location/Qualifiers
source	1. 74

```
/status= Bow 1/0
/db_xref="taxon:10090"
/clone="IMAGE:790392"
/clone_lib="Knowles Solter mouse 2 cell"
/tissue_type="embryo"
```

[illegible][illegible]

RESULT	5			
AI305512/c				
LOCUS	AI305512	82 bp	mRNA	EST
DEFINITION	gq72b02.x1 NCI_CGAP_Ov33	Homo sapiens	CDNA clone	IMAGE:1996587 3'
	mRNA sequence.			

SOURCE ORGANISM	REFERENCE
human.	1 (bases 1 to 82)
Homo sapiens.	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia	
Eutheria; Primates; Catarrhini; Hominoidea; Homo.	

[5']
TGTACGACATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT
3'] double-stranded cDNA was ligated to Eco RI adaptors
[AATTCGATCTTGG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead. "

BASE COUNT	ORIGIN
14 a	36 c 27 g 24 t

Query Match	0.18;	Score 43;	DB 39;	Length 101;
Best Local Similarity	100.0%;	Pred. No. 0.00085;		
Matches 43;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 24577 ATCTCGGCTCACTGCAAGCTCCGGCTCCCGGGTTCACGGCATT 24613
 |||||
 Db 48 ATCTCGGCTCACTGCAAGCTCCGGCTCCCGGGTTCACGGCATT 90

RESULT	8		
LOCUS	AI864984		
DEFINITION	AI864984	90 bp	EST
ACCESSION	U06611.x1	NCI-CGAP_Lym12	Homo sapiens cDNA clone IMAGE:24115
VERSION	AI864984		similar to contatns Alu repetitive element, mRNA sequence.
KEYWORDS	AI864984.1	GI:5529091	
SOURCE	EST.		
ORGANISM	human.		
	Homo sapiens		

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 90)	NCI-CCAP	http://www.ncbi.nlm.nih.gov/ncicgap .		
	National Cancer Institute,	National Cancer Institute, Cancer Genome Anatomy Project (CGAP).		
	Tumor Gene Index			
	Unpublished (1997)			
	On May 18, 1998	this sequence version replaced gi:3137794.		

JOURNAL COMMENT
Unpublished (1997) On May 18, 1998 this sequence version replaced g1:3137794.
Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Life Technologies catalog #: 11547-015
DNA sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html

Seq primer: -40UP from Glibco.	Location/Qualifiers	source
1. .90		

```

/lab_host="DH10B"
/ate="Organ: lymph node; Vector: PCMW-SPORT6; Site:1;
Sali: Site2: Notti; Cloned unidirectionally; Primer:
Oligo dt: Average insert size 1.25 kb. Life Technologies
catalog #: 11547-015"

BASE COUNT      13 a      29 c      27 g      21 t
ORIGIN

```

```

Query Match      0.1%; Score 41; DB 61; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 41; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

Qy 24577 ATCTGGGCTCACTGCAAGTCCGCCCTCCGGGTTACAGCCA 24617
      |||
Db 50 ATCTGGGCTCACTGCAAGTCCGCCCTCCGGGTTACAGCCA 90

```

RESULT	9
AA835411	
LOCUS	AA835411 97 bp mRNA EST 23-FEB-1998
DEFINITION	A772606.s1 Barstead spleen HPLRB2 Homo sapiens cDNA clone IMAGE:113466 3' similar to contains Alu repetitive element; , mRNA sequence.
ACCESSION	AA835411
VERSION	AA835411.1 GI:2909139
KEYWORDS	EST.
SOURCE	human.

REFERENCE	AUTHORS
1 (bases 1 to 97)	
	Hillier, L., Allen, M., Bowles, L., Dubnugue, T., Gatsel, G., Jost, S., Krizman, D., Krcaba, T., Lacy, M., Le, N., Lennon, G., Marrs, M., Martin, J., Moore, B., Scheilenberg, K., Stepcue, M., Tan, F., Theising, B., White, J., Wylie, T., Waterston, R. and Wilson, R.
TITLE	WashU-NCI human EST Project
JOURNAL	Unpublished (1997)
COMMENT	On May 8, 1995 this sequence version replaced g1:801230.

Email: estewatson.mustli.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m3 fwd. ET from Amerham.
infectious/Donal/clone

```

FEATURES
source
    1..97
location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:1413466"
    /clone_id="Barstead spleen HPLRB2"
    /sex="male"
    /dev_stage="adult, 17 years"
    /lab_host="DH10B"
    /note="Organ: spleen; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site.1: EcoRI; Site.2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5
TGTACGCAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTT
3'); double-stranded cDNA was ligated to Eco RI adaptors
[ATTCGATCCTTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead."
BASE COUNT
    16 a
    28 c
    24 g
    29 t
ORIGIN

```

```

Query Match      0.1%: Score 41; DB 39; Length 97;
Best Local Similarity 100.0%: Pred. No. 0.0039;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11894 AGCGTGGAGTGCAGTGGTGCATCTCACTCACTCAAGCT 11934
      |||||
Db 31 AGCGTGGAGTGCAGTGGTGCATCTCACTCACTCAAGCT 71

```

RESULT 10			
LOCUS	AA129957		
DEFINITION	AA129957 104 bp mRNA	EST	27-NOV-1996
IMAGE:565111	5' similar to contains Alu repetitive element;	mRNA	
sequence.	AA129957		
ACCESSION	AA129957.1	GI:1689679	
VERSION	EST.		
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		

REFERENCE 1 (bases 1 to 104)
 AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B., Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, E., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Marais, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Scheilenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevasakis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (9), 807-828 (1996)
 MEDLINE 97044478
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1404721.
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -28m3 rev2 from Amersham
 High quality sequence stop: 92.

FEATURES
 source
 1. .104
 /organism="Homo sapiens"
 /db_xref="GDB:4594175"
 /db_xref="taxon:9606"
 /clone="IMAGE:565111"
 /clone_lib="Stratagene lung carcinoma 937218"
 /tissue_type="lung carcinoma"
 /cell_line="NCI-H69"
 /dev_stage="cell line NCI-H69"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: lung; Vector: pBluescript SK-; Site: 1; EcoRI; Site: 2; XhoI; Cloned unidirectionally. Primer: Oligo dt. Small cell carcinoma cell line NCI-H69. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'."

BASE COUNT 15 a 28 c 26 g 35 t
 ORIGIN

Query Match 0.1%; Score 41; DB 29; Length 104;
 Best Local Similarity 100.0%; Pred. No. 0.0038;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 676 TGATCTGGCTCAGTCACACCTCCGCTCTCTGGGTTCAAG 716
 ||||||||||||||||||||||||||||||||||||||||||||
 DB 50 TGATCTGGCTCAGTCACACCTCCGCTCTCTGGGTTCAAG 90

RESULT 11
 AA669571 95 bp mRNA EST 20-NOV-1997
 LOCUS ac16d10.s1 Stratagene ovary (#937217) Homo sapiens cDNA clone
 DEFINITION IMAGE:856819 3' similar to contains Alu repetitive element; mRNA
 sequence.
 AA669571
 EST. AA669571.1 GI:2631070
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 95)
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Scheilenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE WashU-NCI human EST Project
 JOURNAL Unpublished (1997)
 COMMENT On Nov 6, 1997 this sequence version replaced gi:932677.

CONTACT: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40m3 fwd. ET from Amersham.

FEATURES
 source
 1. .95
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:856819"
 /clone_lib="Stratagene ovary (#937217)"
 /sex="female"
 /dev_stage="49 year old"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="Organ: ovary; Vector: Bluescript SK; Site: 1; EcoRI; Site: 2; XhoI; Cloned unidirectionally. Primer: Oligo dt. Total ovary tissue, normal, caucasian. Average insert size: 0.8 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'."

BASE COUNT 15 a 30 c 24 g 26 t
 ORIGIN

Query Match 0.1%; Score 40; DB 37; Length 95;
 Best Local Similarity 100.0%; Pred. No. 0.0084;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 676 TGATCTGGCTCAGTCACACCTCCGCTCTCTGGGTTCAAG 715
 ||||||||||||||||||||||||||||||||||||||||||||
 DB 30 TGATCTGGCTCAGTCACACCTCCGCTCTCTGGGTTCAAG 69

RESULT 12
 R67088 96 bp mRNA EST 30-MAY-1995
 LOCUS y130h05.s1 Soares placenta Nb2HP Homo sapiens cDNA clone
 DEFINITION IMAGE:140793 3', mRNA sequence.
 R67088
 EST. R67088.1 GI:839726
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 96)
 AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT On Apr 5, 1995 this sequence version replaced gi:760768.
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 1093
 High quality sequence stops: 65
 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert length: 1093 Std Error: 0.00
 Seq primer: Promega -21m3
 High quality sequence stop: 65.
 Location/Qualifiers


```
source
1. .96
/organism="Homo sapiens"
/db_xref="GDB:549405"
/db_xref="taxon:9606"
/clone_lib="IMAGE:140793"
/clone_lib="Soares placenta Nb2Hp"
/sex="Female"
/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: Placenta; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5']
ACTGGAGAAATTCGGCCGCGAGAAATTTTCTTTTCTTTT 3'),
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo. "
```

BASE COUNT 25 a 24 c 21 g 25 t 1 others

Query Match 0.1%; Score 40; DB 22; Length 96;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24749 CCCGCTCGGCTCCCAAGTGTGGATTACAGGCTGA 24788
DB 1 CCCGCTCGGCTCCCAAGTGTGGATTACAGGCTGA 40

RESULT 13
LOCUS AU072757 73 bp mRNA EST 24-JUN-1999
DEFINITION AU072757 Dictyostelium discoidium SS (H.Urushihara) Dictyostelium
discoidium cDNA clone SSA765, mRNA sequence.
ACCESSION AU072757
VERSION AU072757.1 GI:5179178
KEYWORDS EST.
SOURCE Dictyostelium discoidium.
ORGANISM Dictyostelium discoidium
Eukaryota; Dictyostelid; Dictyostelium.
REFERENCE 1 (bases 1 to 73)
AUTHORS Urushihara, H.
JOURNAL Developmental cDNA in Dictyostelium discoidium (1999)
TITLE Unpublished (1999)
CONTACT On Jun 5, 1998 this sequence version replaced gi:3188495.
CONTACT: Hideko Urushihara
Institute of Biological Sciences
University of Tsukuba
3-3-10 Ten-noda, Tsukuba, Ibaraki 305, Japan
Email: d402huesakura.cc.tsukuba.ac.jp
PROJECT = Dictyostelium discoidium cDNA project in Japan.
Location/Qualifiers
1. .73
/organism="Dictyostelium discoidium"
/strain="AX4"
/db_xref="taxon:44689"
/clone_lib="SSA765"
/clone_lib="Dictyostelium discoidium SS (H.Urushihara)"
/dev_stage="slug" 1 g 26 t

BASE COUNT 46 a 0 c 1 g 26 t

ORIGIN

Query Match 0.1%; Score 39; DB 51; Length 73;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10446 AATTAATAATAATAATAATAATAATAATAATAATAAT 10484
DB 15 AATTAATAATAATAATAATAATAATAATAATAATAAT 53.

RESULT 14
LOCUS AI262095/c 80 bp mRNA EST 13-NOV-1998
DEFINITION q228c05.y1 NCI-CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2028200 5'
element /, mRNA sequence.
ACCESSION AI262095.1 GI:3870298
VERSION AI262095.1 GI:3870298
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 80)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 14, 1998 this sequence version replaced gi:1798653.
CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Koskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Seq primer: -40RP from Gibco.
Location/Qualifiers
1. .80
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:2028200"
/clone_lib="NCI-CGAP_Kid11"
/lab_host="DH10B"
/notes="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI-CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 132376-132391, 145607-145675, and
150552-150285). Subtraction by Bento Soares and M.
Fatima Bonaldo. "

BASE COUNT 13 a 25 c 24 g 18 t

ORIGIN

Query Match 0.1%; Score 39; DB 44; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5646 CCACTGCACCTCCAGCTGGGTGACAGCAAGACTCCAT 5644
DB 43 CCACTGCACCTCCAGCTGGGTGACAGCAAGACTCCAT 5

RESULT 15
LOCUS AU038446 93 bp mRNA EST 29-MAR-1999
DEFINITION AU038446 Dictyostelium discoidium SS (H.Urushihara) Dictyostelium
discoidium cDNA clone SSH729, mRNA sequence.
ACCESSION AU038446
VERSION AU038446.1 GI:3985199
KEYWORDS EST.
SOURCE Dictyostelium discoidium.
ORGANISM Dictyostelium discoidium

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 60)
AUTHORS de la Monte,S. and Wands,J.R.
TITLE Neutral thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 57 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..60
BASE COUNT 12 a 14 c 15 g 19 t
ORIGIN

Query Match 0.2%; Score 46; DB 5; Length 60;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25363 GCTAATTTTGTATTTAGTAGACAGAGGTTTCACCGTGTGGC 25408
Db 11 GCTAATTTTGTATTTAGTAGACAGAGGTTTCACCGTGTGGC 56

RESULT 2
LOCUS AF087511 66 bp mRNA PRI 13-SEP-1999
DEFINITION Homo sapiens clone ENAC+22 epithelial sodium channel alpha subunit (SCNNIA) mRNA, alternatively spliced, partial sequence.
ACCESSION AF087511
VERSION AF087511.1 GI:5870626
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 66)
AUTHORS Oh,Y. and Warnock,D.G.
TITLE An Alu cassette in the human epithelial sodium channel
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 66)
AUTHORS Oh,Y. and Warnock,D.G.
TITLE Direct Submision
JOURNAL Submitted (26-AUG-1998) Medicine, UAB, 1720 7th Ave. So., Birmingham, AL 35294, USA
FEATURES Location/Qualifiers
source 1..66
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="ENAC+22"
misc-feature <1..>66
/gene="SCNNIA"
/note="alternatively spliced epithelial sodium channel alpha subunit containing intronic Alu sequence"
gene <1..>66
/gene="SCNNIA"
BASE COUNT 12 a 20 c 24 g 10 t
ORIGIN

Query Match 0.2%; Score 46; DB 40; Length 66;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8738 CTCGTATCCGCCCTCGGCTCCCAAGTCTGGATTACAG 8783
Db 66 CTCGTATCCGCCCTCGGCTCCCAAGTCTGGATTACAG 21

RESULT 3
AR051499

LOCUS AR051499 76 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 69 from patent US 5830670.
ACCESSION AR051499
VERSION AR051499.1 GI:5974863
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 76)
AUTHORS de la Monte,S. and Wands,J.R.
TITLE Neutral thread protein gene expression and detection of Alzheimer's disease
JOURNAL Patent: US 5830670-A 69 03-NOV-1998;
FEATURES Location/Qualifiers
source 1..76
BASE COUNT 15 a 19 c 18 g 24 t
ORIGIN

Query Match 0.2%; Score 46; DB 5; Length 76;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25363 GCTAATTTTGTATTTAGTAGACAGAGGTTTCACCGTGTGGC 25408
Db 11 GCTAATTTTGTATTTAGTAGACAGAGGTTTCACCGTGTGGC 56

RESULT 4
LOCUS HUMBRKFAA 80 bp DNA PRI 27-APR-1993
DEFINITION Human alpha-galactosidase breakpoint region.
ACCESSION M36131
VERSION M36131.1 GI:1179541
KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 80)
AUTHORS Kornreich,R., Bishop,D.F. and Desnick,R.J.
TITLE Alpha-galactosidase A gene rearrangements causing Fabry disease: identification of short direct repeats at breakpoints in an Alu-rich gene
JOURNAL J. Biol. Chem. 265, 9319-9326 (1990)
MEDLINE 90264427
FEATURES Location/Qualifiers
source 1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 19 a 28 c 20 g 13 t
ORIGIN

Query Match 0.2%; Score 45; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19010 CCCACCTTGCCCTCCCAAGTCTGGATTACAGCATGAGCCAC 19054
Db 13 CCCACCTTGCCCTCCCAAGTCTGGATTACAGCATGAGCCAC 57

RESULT 5
LOCUS HUMBRKFAA 80 bp DNA PRI 27-APR-1993
DEFINITION Human alpha-galactosidase breakpoint family F.
ACCESSION M36132
VERSION M36132.1 GI:1179542
KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 80)
Kornreich, R., Bishop, D.F. and Desnick, R.J.
Alpha-galactosidase A gene rearrangements causing Fabry disease:
identification of short direct repeats at breakpoints in an
Alu-rich gene
JOURNAL
J. Biol. Chem. 265, 9319-9326 (1990)
MEDLINE
90264427
FEATURES
source
Location/Qualifiers
1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"

Query Match
Best Local Similarity 100.0%; Score 45; DB 9; Length 80;
Pred. No. 3.2e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19010 CCCACCTGGCTGCCAAGTCTGGATACAGCATGAGCCAC 19054
|||||
Db 13 CCCACCTGGCTGCCAAGTCTGGATACAGCATGAGCCAC 57
|||||

RESULT 6
HUMUT578A/c
LOCUS HUMUT578A 69 bp DNA STS 19-JUL-1995
DEFINITION Human STS UT578, 5' primer bind, sequence tagged site.
ACCESSION L39139 L18336
VERSION L39139.1 GI:642101
KEYWORDS STS; PCR primer; STS sequence; microsatellite DNA; microsatellite
marker; sequence tagged site; tetranucleotide repeat.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 69)
Gerken, S.C., Matsunami, N., Plaetke, R., Albertsen, H., Ballard, L.,
Wells, R., Lawrence, E., Moore, M., Holik, P.R., Carlson, M., Zhao, X.,
Robertson, M., Bradley, P., Elsnier, T., Tingey, A., Lalouel, J.-M. and
White, R.
Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
Unpublished (1994)
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: stecorona.med.utah.edu
Primer A: AGTGGAGACCGCTGGC
Primer B: AGTGGAGACCGCTGGC
End to Label: Primer A
PCR Profile:
Initial Denaturation: 94C 30sec
Cycles Denaturation Annealing Extension 5 94
C 10 sec. 62 C 10 sec. 72 C 20 sec. 30
58 C 10 sec. 72 C 20 sec. Mg++: 1.50 mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 1.
FEATURES
source
Location/Qualifiers
1..69
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="19"
14..32
/note="STS UT578 5' end"
/evidence="experimental"
BASE COUNT 26 a 16 c 14 g 13 t

Query Match
Best Local Similarity 100.0%; Score 44; DB 13; Length 69;
Pred. No. 1e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21714 TTCACATGTTGGCAGGCTGTCTGCAACTCTGACCTCACT 21757
|||||
Db 44 TTCACATGTTGGCAGGCTGTCTGCAACTCTGACCTCACT 1
|||||

RESULT 7
HUMALUANC/c
LOCUS HUMALUANC 65 bp DNA PRI 08-OCT-1994
DEFINITION Homo sapiens 4000 year old remains from Nekht-ankh Alu repeat 12:2.
ACCESSION L36836
VERSION L36836.1 GI:556194
KEYWORDS Homo sapiens (individual isolate 4000 year old remains from
Nekht-ankh) liver DNA.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (sites)
Paabo, S.
Ancient DNA: extraction, characterization, molecular cloning, and
enzymatic amplification
Proc. Natl. Acad. Sci. U.S.A. 86 (6), 1939-1943 (1989)
MEDLINE
89184542
FEATURES
source
Location/Qualifiers
1..65
/organism="Homo sapiens"
/isolate="4000 year old remains from Nekht-ankh"
/db_xref="taxon:9606"
/tissue_type="liver"
repeat_region
17 a 17 c 21 g 10 t

Query Match
Best Local Similarity 100.0%; Score 43; DB 9; Length 65;
Pred. No. 3.3e-10;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8735 GACCTGTGATCGCCGCTGCTGCCAAGTCTGGAT 8777
|||||
Db 65 GACCTGTGATCGCCGCTGCTGCCAAGTCTGGAT 23
|||||

RESULT 8
HUMALUANC/c
LOCUS HUMALUANC 67 bp DNA PRI 08-OCT-1994
DEFINITION Homo sapiens 4000 year old remains from Nekht-ankh Alu repeat
fragment 12:5.
ACCESSION L36843
VERSION L36843.1 GI:556196
KEYWORDS Alu repeat.
SOURCE Homo sapiens (individual isolate 4000 year old remains from
Nekht-ankh) liver DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (sites)
Paabo, S.
Ancient DNA: extraction, characterization, molecular cloning, and
enzymatic amplification
Proc. Natl. Acad. Sci. U.S.A. 86 (6), 1939-1943 (1989)
MEDLINE
89184542
FEATURES
source
Location/Qualifiers
1..67
/organism="Homo sapiens"
/isolate="4000 year old remains from Nekht-ankh"
/db_xref="taxon:9606"
/tissue_type="liver"

repeat_region <1..>67
/rpt_family="Alu"
BASE COUNT 15 a 18 c 23 g 11 t
ORIGIN

Query Match 0.1%; Score 43; DB 9; Length 67;
Best Local Similarity 100.0%; Pred. No. 3.3e-10;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25448 GCGTCAGCCTCCCAAGTGTGGATTACAGGTGAGCCACC 25490
|||||
Db 50 GCGTCAGCCTCCCAAGTGTGGATTACAGGTGAGCCACC 8

RESULT 9
A68621/c A68621 40 bp DNA PAT 06-MAY-1999
LOCUS Sequence 1 from Patent WO9801573.
DEFINITION A68621
ACCESSION A68621
VERSION A68621.1 GI:4759648
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 40)
AUTHORS Resnick,M.A., Lariouov,V.L., Koudrina,N.Y. and Perkins,E.L.
TITLE TRANSFORMATION-ASSOCIATED RECOMBINATION CLONING
JOURNAL Patent: WO 9801573-A 15-JAN-1998;
US HEALTH (US)

FEATURES
Source Location/Qualifiers
1..40
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 7 a 12 c 13 g 8 t
ORIGIN

Query Match 0.1%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8755 TCGGCTCCCAAGTGTGGATTACAGCGTGAGCCACC 8794
|||||
Db 40 TCGGCTCCCAAGTGTGGATTACAGCGTGAGCCACC 1

RESULT 10
HUMALCE221/c HUMALCE221 103 bp ss-RNA PRI 15-APR-1994
LOCUS Human carcinoma cell-derived Alu RNA transcript, clone CE221.
DEFINITION M87896
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 103)
AUTHORS Slonett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press

FEATURES
Source Location/Qualifiers
1..103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="NTera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"

BASE COUNT 25 a 27 c 33 g 18 t
ORIGIN

Query Match 0.1%; Score 40; DB 9; Length 103;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18828 CGATCTGGCTCACTGCACCTCCGCTCCGGGTTCAAG 18867
|||||
Db 87 CGATCTGGCTCACTGCACCTCCGCTCCGGGTTCAAG 48

RESULT 11
AR051522 AR051522 85 bp DNA PAT 29-SEP-1999
LOCUS Sequence 92 from patent US 5830670.
DEFINITION AR051522
ACCESSION AR051522
VERSION AR051522.1 GI:5974886
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 85)
AUTHORS de la Monte,S. and Wands,J.R.
TITLE Neural thread protein gene expression and detection of Alzheimer's
disease
JOURNAL Patent: US 5830670-A 92 03-NOV-1998;
US HEALTH (US)

FEATURES
Source Location/Qualifiers
1..85
/organism="unknown"
BASE COUNT 13 a 27 c 25 g 20 t
ORIGIN

Query Match 0.1%; Score 39; DB 5; Length 85;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17462 GCCAGATGCTCGATCTCCTGACCTCGTGATCCGCC 17500
|||||
Db 9 GCCAGATGCTCGATCTCCTGACCTCGTGATCCGCC 47

RESULT 12
HUMBRKFAC HUMBRKFAC 80 bp DNA PRI 27-APR-1993
LOCUS Human alpha-galactosidase breakpoint region.
DEFINITION M36133
ACCESSION M36133
VERSION M36133.1 GI:179543
KEYWORDS Fabry disease; breakpoint junction; glycosphingolipid catabolism.
SOURCE Human DNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 80)
AUTHORS Kornreich,R., Bishop,D.F. and Desnick,R.J.
TITLE Alpha-galactosidase A gene rearrangements causing Fabry disease:
Identification of short direct repeats at breakpoints in an
Alu-rich gene

JOURNAL J. Biol. Chem. 265, 9319-9326 (1990)
MEDLINE 90264427
FEATURES
Source Location/Qualifiers
1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 18 a 28 c 17 g 17 t
ORIGIN

Query Match 0.1%; Score 37; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19018 GGCTCCCAAGTGTGGATTACAGGATGAGCCAC 19054

```
|||||
Db      21 GGCCTCCCAAAGTCTGGGATTACAGCATGAGCCAC 57
RESULT  13
LOCUS   HUMDLRFL 90 bp DNA PRI 11-JAN-1995
DEFINITION Human low density lipoprotein receptor Intron A Alu repeat.
ACCESSION K03555.1 GI:187104
VERSION   K03555.1 GI:187104
KEYWORDS  Alu repeat; low density lipoprotein receptor-1; repeat region.
SOURCE    Human DNA.
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 90)
AUTHORS   Eutheria; Primates; Carnivora; Mammalia;
           Lehman,M.A., Goldstein,J.L., Russell,D.W. and Brown,M.S.
           Duplication of seven exons in LDL receptor gene caused by Alu-Alu
           recombination in a subject with familial hypercholesterolemia
           Cell 48 (5), 827-835 (1987)
JOURNAL   87131094
COMMENT   Clean copy of sequence kindly provided by M. Lehman (22-APR-1987).
FEATURES  Location/Qualifiers
           source          1..90
                           /organism="Homo sapiens"
                           /db_xref="taxon:9606"
                           /map="19p13.3"
           misc_feature     1
                           /gene="LDLR"
                           /note="hgml: G00-119-362"
                           /map="19p13.3"
BASE COUNT 17 a 30 c 24 g 19 t
ORIGIN      Chromosome 19p13.2-p13.1.

Query Match 0.1%; Score 37; DB 9; Length 90;
Best Local Similarity 100.0%; Pred. No. 3.4e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19018 GGCCTCCCAAAGTCTGGGATTACAGCATGAGCCAC 19054
|||||
Db      42 GGCCTCCCAAAGTCTGGGATTACAGCATGAGCCAC 78
RESULT  14
LOCUS   I31263/c 40 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 175 from patent US 5582979.
ACCESSION I31263
VERSION   I31263.1 GI:1822054
KEYWORDS
SOURCE    Unknown.
ORGANISM  Unclassified.
REFERENCE 1 (bases 1 to 40)
AUTHORS   Weber,J.L.
TITLE     Length polymorphisms in (dc-da).sub.n.(dc-dt).sub.n sequences and
           method of using the same
JOURNAL   Patent: US 5582979-A 175 10-DEC-1996;
           Location/Qualifiers
           source          1..40
                           /organism="unknown"
BASE COUNT 20 a 19 c 0 g 1 t
ORIGIN

Query Match 0.1%; Score 36; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28999 TGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 29034
|||||
Db      36 TGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1
```

```
|||||
RESULT  15
LOCUS   I31159/c 46 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 71 from patent US 5582979.
ACCESSION I31159
VERSION   I31159.1 GI:1821950
KEYWORDS
SOURCE    Unknown.
ORGANISM  Unclassified.
REFERENCE 1 (bases 1 to 46)
AUTHORS   Weber,J.L.
TITLE     Length polymorphisms in (dc-da).sub.n.(dc-dt).sub.n sequences and
           method of using the same
JOURNAL   Patent: US 5582979-A 71 10-DEC-1996;
           Location/Qualifiers
           source          1..46
                           /organism="unknown"
BASE COUNT 23 a 22 c 0 g 1 t
ORIGIN

Query Match 0.1%; Score 36; DB 5; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28999 TGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 29034
|||||
Db      44 TGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 9
```

Search completed: April 1, 2000, 10:12:16
Job time: 283685 sec

Mon Apr 3 08:24:19 2000

us-08-852-495c-1_copy_206000_235033.rge

Page 6

PR 21-APR-1989; US-341562.
 PR 05-SEP-1991; US-754351.
 PR 04-APR-1994; US-222177.
 PA (MARS-) MARSHFIELD CLINIC.
 PI Weber, J.L.
 DR WPI; 97-042299/04.
 PT Detection of polymorphic genetic markers of the form
 PT (dc-da)n(dc-dr)n - using novel nucleic acid mols. as primers
 PS Disclosure; Column 9-10; 186pp; English.
 CC The invention relates to the isolation of polymorphic repeat sequences
 CC having the sequence (dc-da)n.(dc-dr)n which can be used as genetic
 CC markers. Primers based on these sequences can be used to detect these
 CC repeats, especially for use in e.g. paternity or maternity testing,
 CC human genetic analysis such as linkage analysis of genetic disease,
 CC commercial animal or plant breeding or pedigree analysis. Clones
 CC containing the repeat sequences were isolated by hybridisation of
 CC chromosome-specific phage libraries with a synthetic poly(dc-da).(dc-dr)
 CC probe. Over 100 repeat blocks were isolated. The inserts from the clones
 CC were amplified by primers T65798-T66047. Those clones where the repeat
 CC sequence has been determined are shown in T65704-T797. This repeat
 CC sequence is from the marker clone Mdf42 which contains the repeat
 CC sequence having the formula: (CA)16T(AC)3.5.
 CC Sequence 40 Bp, 20 A, 19 C, 0 G, 1 T,

[illegible]

RESULT 3
ID T65709/c
AC T65709 standard; DNA: 46 BP.
DT 17-JUN-1997 (first entry)
DE Repeat sequence from polymorphic marker clone Mfd7.
KW Polymorphism; repeat sequence; genetic marker; primer; amplification;
PCR; polymerase chain reaction; paternity; maternity; human; pedigree;
linkage analysis; genetic disease; animal; plant; breeding; locus;
hybridisation; chromosome; ds.
OS Homo sapiens.
PN US5582979-A.
PD 10-DEC-1996.
PE 21-APR-1989; 341562.
PR 21-APR-1989; US-341562.
PR 05-SEP-1991; US-754351.
PR 04-APR-1994; US-222177.
PA (MARS-) MARSHFIELD CLINIC.
PI Weber JL;
DR WPI: 97-042299/04.
PT Detection of polymorphic genetic markers of the form
PT (dc-da)n(dg-dt)n - using novel nucleic acid mols. as primers
PS Claim 1; Column 9-10; 186pp; English.
CC The invention relates to the isolation of polymorphic repeat sequences
CC having the sequence (dc-da)n.(dg-dt)n which can be used as genetic
CC markers. Primers based on these sequences can be used to detect these
CC repeats, especially for use in e.g. paternity or maternity testing,
CC human genetic analysis such as linkage analysis of genetic disease,
CC commercial animal or plant breeding or pedigree analysis. Clones
CC containing the repeat sequences were isolated by hybridisation of
CC chromosome-specific phage libraries with a synthetic poly(dc-da) (dg-dt)
CC probe. Over 100 repeat blocks were isolated. The inserts from the clones
CC were amplified by primers T65798-T66047. Those clones where the repeat
CC sequence has been determined are shown in T65704-T67197. This repeat
CC sequence is from the marker clone Mdf7 which contains the repeat
CC sequence having the formula: (CA)20TA(CA)2.
SQ Sequence 46 BP; 23 A; 22 C; 0 G; 1 T;

Query Match	0.1%	Score 36	DB 1	Length 46
Best Local Similarity		100.0%		
Matches	36	Conservative	0	Mismatches
			0	Indels
			0	Gaps

```
QY    28999 TGTATGCTGCTGTCGTCGTGTCGTGTCGTG 29034
```

```
Dd     44 TGTATGCTGCTGTCGTCGTGTCGTGTCGTG 9
```

RESULT 4

AC Q34122;

Sequence of a microsatellite from clone WGLA68
PCR: selection: primers: OPTIPRIM: breeding: C

Bos taurus.

PD 00-AUG-1992;
PF 15-JAN-1992; U00340.

PA (GEM-) GENMARK.
PT Georges M. Massey TM:

PT Polymorphic bovine DNA markers - used in genetic identification,

PS Table 1, page 3/5, 21/1pp, English.

CC The sequence is that of a bovine microsatellite sequence obtd.

CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
CC One cut of 50 clones cross-hybridised. Assuming independent

CC (T6)n >9 microsatellites in the bovine genome is estimated at >100,

CC The sequences upstream and downstream of the microsatellite sequence
CC is summarised in the specification and indexed herein (see below).

CC amplification of the corresp. microsatellite (using the program

CC for percentage testing, and in the genetic mapping of economically important
CC loci, or genes involved the determinism of economically important

CC See also Q33501-34437.
Sequence 49 Rp: 7 A: 0 C: 17 G: 25 T: 50

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

Best Local Samtality: 100.00; FRC: 20.20;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
Qy      28999 TGTATGTCGTCTGTCTGTCTGTCTGTCTGTCTG 29033
          |||||
Db       1 TGTATGTCGTCTGTCTGTCTGTCTGTCTGTCTG 36
```

RESUL 3
Q33870

AC Q33870; DE 03-FEB-1963 /E4 50+ 00+ 50+ 50+

PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage; KW

OS Bos taurus.
PN W09213102-A

PF 15-JAN-1992; U00340.

PA (GENM-) GENMARK.
PI Georges M. Massey JM:

PT Polymorphic bovine DNA markers - used in genetic identification,

PT gene mapping, and selective breeding
 PS Table 7; Page 277; 517bp; English.
 CC The sequence is that of a bovine microsatellite sequence obtd. by
 CC screening a library of bovine MboI DNA fragments of between
 CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
 CC One out of 50 clones cross-hybridised. Assuming independent
 CC distribution of microsatellites and MboI sites, the frequency of
 CC (T6)n >9 microsatellites in the bovine genome is estimated at >100.
 CC 000. The sequence information for ca. 230 such bovine microsatellites
 CC is summarised in the specification and indexed herein (see below).
 CC The sequences upstream and downstream of the microsatellite sequence
 CC were used to generate the required PCR primers for in vitro
 CC amplification of the corresp. microsatellite (using the program
 CC OPR1PRM). The microsatellites may be used to identify individuals,
 CC for parentage testing, and in the genetic mapping of economic trait
 CC loci, or genes involved in the determinism of economically important
 CC traits esp. in cattle, to allow selective breeding.
 CC See also Q33501-34437.
 SO Sequence 54 Bp; 1 A; 0 C; 26 G; 27 T;

Query Match	0.18;	Score 36;	DB 1;	Length 54;
Best Local Similarity	100.0%;	Pred. No. 2.5;		
Matches 36;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	28939	TCATGCTGCTGCTGCTGCTGCTGCTGCTG	29034
DB	11	TCATGCTGCTGCTGCTGCTGCTGCTGCTG	46

RESULT	6
Q27391	
ID	Q27391 standard; DNA; 35 BP.

DT 27-JAN-1993 (first entry) PDJ33.
DE Inter-Alu specific primer
KW Polymerase chain reaction; PCR; repetitive element; ss.
OS Synthetic.
PN WO9213101-A.
PD 06-AUG-1992.
PF 24-JAN-1992; NL00018.
PR 25-JAN-1991; NL-000132.
PA (INGE-) INGENY BV.
PI Uiterlinden AG, V1jg J;
DR WPI: 92-284683/34.
PT Detection of genetic variation by 2-D electrophoresis of
PT fragments - and hybridisation with labelled probes; carried out
PT on fragments consisting of inter-repeat sequences generated by
PT PCR.
PS Claim 6; Page 6; 31pp; English.
PS Primer PDJ33 is one of several primers which are preferred for use
CC in amplifying inter-Alu regions of DNA. The amplified fragments are
CC then subjected to 2-D electrophoresis on the basis of length and
CC differences in base sequence. The resulting separation pattern is
CC transferred to a filter for screening with a probe. The method can
CC be used to detect genetic variation.
CC See Q27385-Q27404 and Q33141-Q33144.
SO Sequence 35 BP; 8 A; 10 C; 11 G; 6 T;

Query Match	0.1%;	Score 35;	DB 1;	Length 35;
Best Local Similarity	100.0%;	Pred. No. 4.9;		
Matches	35;	Conservative	0;	Mismatches
			0;	Indels
			0;	Gaps
			0;	

DQ 8758 GCGTCCCAAGTGCTGGATTCACGCGGTGAGCCA 8792
|||||
|||||
Db 1 GCGTCCCAAGTGCTGGATTACAGCGGTGAGCCA 35
|||||
|||||

RESULT 7
T65737/c
ID T65737 standard; DNA; 45 BP.
AC T65737;

DE 17-JUN-1997 (first entry)
DE Repeat sequence from polymorphic marker clone Mf336.
KM Polymorphism; repeat sequence; genetic marker; primer; amplification;
KM PCR; polymerase chain reaction; paternity; maternity; human; pedigree;
KM linkage analysis; genetic disease; animal; plant; breeding; locus;
KM hybridisation; chromosome; ds.

PN US5582979-A.
PD 10-DEC-1996.
PF 21-APR-1989. 341562.
PR 21-APR-1989. US-541562.
PR 05-SEP-1991. US-754351.
PR 04-APR-1994. US-222177.
PA (MARS-) MARSHFIELD CLINIC.

PT Detection of polymorphic genetic markers of the form
 (dc-da)n(dg-dt)n - using novel nucleic acid mols. as primers
 PS Disclosure; Column 9-10; 186pp; English.

CC markers. Primers based on these sequences can be used to detect these
CC repeats, especially for use in e.g. paternity or maternity testing,
CC human genetic analysis such as linkage analysis of genetic disease,

CC chromosome-specific phage libraries with a synthetic poly(dC-dA). (dG-dT)
CC probe. Over 100 repeat blocks were isolated. The inserts from the clones
CC were amplified by primers T65798-rT66047. Those clones where the repeat
CC sequence has been determined are shown in T65704-797. This repeat
CC sequence is from the marker clone Kd136 which contains the repeat
CC sequence having the formula: (AC)15A(AC)6A.
SQ Sequence 45 BP; 21 C; 0 G; 1 T;

Query Match	0.1%	Score 35	DB 1	Length 45
Best Local Similarity	100.0%	Pred. No. 4.4		
Matches 35, Conservative	0	Mismatches	0	Gaps 0

[illegible]

RESULT 8
Q34053/c
ID Q34053 standard; DNA: 62 BP.

DE Microsatellite sequence from clone TGLA438.
 PCR: selection; primers; OpiPrIM; breeding; cattle; parentage;
 KW genetic mapping; traits; amplification; ss.

PD 06-AUG-1992. U00340.
 PE 15-JAN-1992;
 PR 15-JAN-1991; US-642342.

DR WPI: 92-284684/34.
PT Polymorphic bovine DNA markers - used in genetic identification, gene mapping, and selective breeding
Table 7: page 352: 517nn: English
p5

CC screening a library of bovine MboI DNA ligations, between
CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MboI sites, the frequency of

CC is summarised in the specification and indexed herein (see below).
 CC The sequences upstream and downstream of the microsatellite sequence
 CC were used to generate the required PCR primers for *in vitro*

AC	T26728;
DT	22-OCT-1996 (first entry)
DE	Human gene signature HUMGS08978.

CC were used to generate the required PCR primers for *in vitro*

CC amplification of the corresp. microsatellite (using the program
 CC OPIPRIM). The microsatellites may be used to identify individuals,
 CC for parentage testing, and in the genetic mapping of economic trait
 CC loci, or genes involved in the determination of economically important
 CC traits esp. in cattle, to allow selective breeding.
 CC See also Q33501-34437.
 SQ Sequence 44 BP; 22 A; 0 C; 0 G; 22 T;

Query Match 0.1%; Score 33; DB 1; Length 44;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13052 ATATATATATATATATATATATATATATA 13084
 |||||
 DB 3 ATATATATATATATATATATATATATATA 35

RESULT 15

ID Q33855/c
 AC Q33855. standard; DNA; 44 BP.

DT 02-FEB-1993 (first entry)
 DE Microsatellite sequence from clone TGLA261.
 KW PCR; selection; primers; OPIPRIM; breeding; cattle; parentage;
 KW genetic mapping; traits; amplification; ss.
 OS Bos taurus.

PN W09213102-A.

PD 06-AUG-1992.

PF 15-JAN-1992; U00340.

PR 15-JAN-1991; US-642342.

PA (GENM-) GENMARK.

PI Georges M. Massey JM:

DR WPI: 92-284684/34.

PT Polymorphic bovine DNA markers - used in genetic identification,

PT gene mapping, and selective breeding

PS Table 7; Page 271; 517pp; English.

CC The sequence is that of a bovine microsatellite sequence obtd. by

CC screening a library of bovine MboI DNA fragments of between

CC 250 and 500 bp with an (AC)₁₅ and a (TC)₁₅ oligonucleotide probe.

CC One out of 50 clones cross-hybridised. Assuming independent

CC distribution of microsatellites and MboI sites, the frequency of

CC (T6)_n >9 microsatellites in the bovine genome is estimated at >100,

CC 000. The sequence information for ca. 230 such bovine microsatellites

CC is summarised in the specification and indexed herein (see below).

CC The sequences upstream and downstream of the microsatellite sequence

CC were used to generate the required PCR primers for in vitro

CC amplification of the corresp. microsatellite (using the program

CC OPIPRIM). The microsatellites may be used to identify individuals,

CC for parentage testing, and in the genetic mapping of economic trait

CC loci, or genes involved in the determination of economically important

CC traits esp. in cattle, to allow selective breeding.

CC See also Q33501-34437.

SQ Sequence 44 BP; 22 A; 0 C; 0 G; 22 T;

Query Match 0.1%; Score 33; DB 1; Length 44;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13052 ATATATATATATATATATATATATATATA 13084
 |||||
 DB 42 ATATATATATATATATATATATATATATA 10

Search completed: April 1, 2000, 12:51:40
 Job time: 291698 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: April 1, 2000, 04:07:52 ; Search time 329.34 Seconds
(without alignment)

10554.569 Million cell updates/sec

Title: US-08-852-495C-1_COPY_206000_235033
 Document: 20034

Sequence: 1 AGTTGTAATAGTAGACAA.....TGTGTGTGTGTGTGTGTG 29034

Scoring table:	OLIGO_NUC	Count
CCCC	50	0

Searched: 214294 seqs, 59861574 residues

Word size :

Total number of hits satisfying chosen parameters:

349150

```
Minimum DB seq length: 8
Maximum DB seq length: 105
```

Post-processing: Listing first 45 summaries

Database ::

```

1: /cgn2_6/prodata/1/lna/5A_COMB.seq:*
2: /cgn2_6/prodata/1/lna/5B_COMB.seq:*
3: /cgn2_6/prodata/1/lna/5C_COMB.seq:*
4: /cgn2_6/prodata/1/lna/5D_COMB.seq:*
5: /cgn2_6/prodata/1/lna/5_COMB.seq:*
6: /cgn2_6/prodata/1/lna/PCrUS9_COMB.seq:*
7: /cgn2_6/prodata/1/lna/Backfiles1.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	46	0.2	60	3	US-08-454-557C-57	Sequence 57, App
2	46	0.2	60	4	US-08-340-426D-57	Sequence 57, App
3	46	0.2	60	4	US-08-450-673C-57	Sequence 57, App
4	46	0.2	60	6	PCT-US95-17111A-57	Sequence 57, App
5	46	0.2	76	3	US-08-454-557C-69	Sequence 69, App
6	46	0.2	76	4	US-08-340-426D-69	Sequence 69, App
7	46	0.2	76	4	US-08-450-673C-69	Sequence 69, App
8	46	0.2	76	6	PCT-US95-17111A-69	Sequence 69, App
9	39	0.1	85	3	US-08-454-557C-92	Sequence 92, App
10	39	0.1	85	4	US-08-340-426D-92	Sequence 92, App
11	39	0.1	85	4	US-08-450-673C-92	Sequence 92, App
12	39	0.1	85	6	PCT-US95-17111A-92	Sequence 92, App
13	36	0.1	40	1	US-08-222-177A-175	Sequence 175, App
14	36	0.1	46	1	US-08-222-177A-71	Sequence 71, App
15	36	0.1	54	2	US-08-469-802B-27	Sequence 27, App
16	36	0.1	54	3	US-08-267-803B-45	Sequence 45, App
17	35	0.1	45	1	US-08-222-177A-157	Sequence 157, App
18	34	0.1	47	1	US-08-222-177A-458	Sequence 458, App
19	33	0.1	56	1	US-08-222-177A-149	Sequence 149, App
20	33	0.1	84	3	US-08-454-557C-91	Sequence 91, App
21	33	0.1	84	4	US-08-340-426D-91	Sequence 91, App
22	33	0.1	84	4	US-08-450-673C-91	Sequence 91, App
23	33	0.1	84	6	PCT-US95-17111A-91	Sequence 91, App
24	32	0.1	33	1	US-08-222-177A-59	Sequence 59, App
25	32	0.1	33	1	US-08-222-177A-218	Sequence 218, App
26	32	0.1	34	1	US-08-222-177A-110	Sequence 110, App

C 27	32	0.1	34	1	US-08-222-177A-172	Sequence 172, App
C 28	32	0.1	34	1	US-08-222-177A-180	Sequence 180, App
C 29	32	0.1	34	1	US-08-222-177A-192	Sequence 192, App
C 30	32	0.1	34	1	US-08-222-177A-322	Sequence 322, App
C 31	32	0.1	35	1	US-08-222-177A-77	Sequence 77, App
C 32	32	0.1	35	1	US-08-222-177A-169	Sequence 169, App
C 33	32	0.1	36	1	US-08-222-177A-104	Sequence 104, App
C 34	32	0.1	36	1	US-08-222-177A-358	Sequence 358, App
C 35	32	0.1	37	1	US-08-222-177A-140	Sequence 140, App
C 36	32	0.1	38	1	US-08-222-177A-198	Sequence 198, App
C 37	32	0.1	38	1	US-08-222-177A-397	Sequence 397, App
C 38	32	0.1	39	1	US-08-222-177A-137	Sequence 137, App
C 39	32	0.1	40	1	US-08-222-177A-119	Sequence 119, App
C 40	32	0.1	40	1	US-08-222-177A-400	Sequence 400, App
C 41	32	0.1	40	1	US-08-222-177A-403	Sequence 403, App
C 42	32	0.1	41	1	US-08-222-177A-74	Sequence 74, App
C 43	32	0.1	41	1	US-08-222-177A-183	Sequence 183, App
C 44	32	0.1	41	1	US-08-222-177A-232	Sequence 232, App
C 45	32	0.1	42	1	US-08-222-177A-53	Sequence 53, App

ALIGNMENTS

RESULT 1

; Sequence 57, Application US/08454557C
; Patent No. 5830670

APPLICANT: de la

TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

```

;
NUMBER OF SEQUENCES: 121

```

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
ATTENTION: 100 West Street
CITY: New York City

CITY: Washington

COUNTRY: U.S.A

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25

APPLICATION NUMBER: US/08/454, 557C

CLASSIFICATION: 514

NAME: Ludwig, Steven R.

REFERENCE/DOCKET NUMBER: 06

TELEPHONE: (202) 371-2600

; INFORMATION FOR SEQ ID NO: 57:

LENGTH: 60 base pairs

STRANDEDNESS: both

US-08-454-557C-57

Query Match	Score	DB	Length
0.28;	46;	DB 3;	Length 60

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25363 GCTAATTTTGTATTTTTAGTAGAGACAGGGTTTCACCGTGGC 25400

Db 11 GCTATTTTTGTATTTTTAGTAGAGACAGGTTTCACCGTTGGC 56

RESULT 2
US-08-340-426D-57
Sequence 57, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-57
Query Match 0.2%; Score 46; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 6.2e-06;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25363 GCTAATTTTGTATTTTGTAGACAGAGGTTTCACCGTGTGGC 25408
DB 11 GCTAATTTTGTATTTTGTAGACAGAGGTTTCACCGTGTGGC 56
RESULT 3
US-08-450-673C-57
Sequence 57, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-57
Query Match 0.2%; Score 46; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 6.2e-06;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25363 GCTAATTTTGTATTTTGTAGACAGAGGTTTCACCGTGTGGC 25408
DB 11 GCTAATTTTGTATTTTGTAGACAGAGGTTTCACCGTGTGGC 56
RESULT 4
PCT-US95-17111A-57
Sequence 57, Application PC/TUS951711A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and
TITLE OF INVENTION: Detection of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-57

Query Match 0.2%; Score 46; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 6.2e-06;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25363 GCTAATTTTGTATTTTGTAGACAGACGGGTTTCACCGTGTGGC 25408
|||||

Db 11 GCTAATTTTGTATTTTGTAGACAGACGGGTTTCACCGTGTGGC 56

RESULT 5
US-08-454-557C-69
; Sequence 69, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; City: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-454-557C-69

Query Match 0.2%; Score 46; DB 3; Length 76;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25363 GCTAATTTTGTATTTTGTAGACAGACGGGTTTCACCGTGTGGC 25408
|||||

Db 11 GCTAATTTTGTATTTTGTAGACAGACGGGTTTCACCGTGTGGC 56

RESULT 6
US-08-340-426D-69
; Sequence 69, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
City: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-340-426D-69

Query Match 0.2%; Score 46; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25363 GCTAATTTTGTATTTTGTAGACAGACGGGTTTCACCGTGTGGC 25408
|||||

Db 11 GCTAATTTTGTATTTTGTAGACAGACGGGTTTCACCGTGTGGC 56

RESULT 7
US-08-450-673C-69
; Sequence 69, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; City: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-69

Query Match 0.2% Score 46; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25363 GCTAATTTTGTATTTTGTAGACAGACGGTTCCACCGTGTGGC 25408
|||||
DB 11 GCTAATTTTGTATTTTGTAGACAGACGGTTCCACCGTGTGGC 56

RESULT 8
PCT-US95-17111A-69
Sequence 69, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and
TITLE OF INVENTION: Detection of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
CLASSIFICATION:
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2500
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 76 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-69

Query Match 0.2% Score 46; DB 6; Length 76;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25363 GCTAATTTTGTATTTTGTAGACAGACGGTTCCACCGTGTGGC 25408
|||||
DB 11 GCTAATTTTGTATTTTGTAGACAGACGGTTCCACCGTGTGGC 56

RESULT 9
US-08-454-557C-92

Sequence 92, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2500
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-92

Query Match 0.1% Score 39; DB 3; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17462 GCCAGATGCTGTCGATCTCGACCTCGTATCGGCC 17500
|||||
DB 9 GCCAGATGCTGTCGATCTCGACCTCGTATCGGCC 47

RESULT 10
US-08-340-426D-92
Sequence 92, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D

FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-92

Query Match 0.1%; Score 39; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17462 GCCAGATGGTCTGATCTCTGACCTGATCGGCC 17500
|||||
DB 9 GCCAGATGGTCTGATCTCTGACCTGATCGGCC 47

RESULT 11
US-08-450-673C-92
Sequence 92, Application US/08450673C
Patent No. 594888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450.673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-92

Query Match 0.1%; Score 39; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17462 GCCAGATGGTCTGATCTCTGACCTGATCGGCC 17500
|||||
DB 9 GCCAGATGGTCTGATCTCTGACCTGATCGGCC 47

RESULT 12
PCT-US95-17111A-92
Sequence 92, Application PC/TUS9517111A
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17111A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/340,426
FILING DATE: 14-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 85 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
PCT-US95-17111A-92

Query Match 0.1%; Score 39; DB 6; Length 85;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17462 GCCAGATGGTCTGATCTCTGACCTGATCGGCC 17500
|||||
DB 9 GCCAGATGGTCTGATCTCTGACCTGATCGGCC 47

RESULT 13
US-08-222-177A-175/C
Sequence 175, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
NUMBER OF SEQUENCES: (dc-da)n.(dg-dt)n SEQUENCES AND METHODS OF USING SAME
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA

```

ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Saira, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865, 601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TEXT:
INFORMATION FOR SEQ ID NO: 175:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: mfd42rs
US-08-222-177A-175

Query Match          0.1%; Score 36; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 28999 TGTATGTCGCTGTCGTGTCGTGTCGTGTCGTGTCGTGC 29034
Dn      36 TGTATGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCG 1

RESULT 14
US-08-222-177A-71/c
Sequence 71, Application US/08222177A
Patent No. 5582979
GENERAL INFORMATION:
APPLICANT: Weber, James L.
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN
TITLE OF INVENTION: (dc-da)n.(dc-dt)n SEQUENCES AND METHODS OF USING SAME
NUMBER OF SEQUENCES: 460
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dewitt Ross & Stevens, S.C.
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Saira, Charles S.
REGISTRATION NUMBER: 30,492

```

[illegible]

Mon Apr 3 08:24:20 2000

us-08-852-495c-1_copy_206000_235033.rni

Page 7

[illegible]

Search completed: April 1, 2000, 12:41:53
Job time: 291175 sec

C	2	59	0.2	103	104	A0535244
C	3	52	0.2	72	62	A1914923
C	4	48	0.2	102	81	B62983
C	5	48	0.2	103	81	B46914
C	6	47	0.2	91	50	F29274
C	7	47	0.2	105	21	T94466
C	8	44	0.2	70	25	N84707
C	9	44	0.2	80	50	F24490
C	10	43	0.1	87	20	D20989
C	11	43	0.1	101	39	A4835205
C	12	41	0.1	90	61	A1864984
C	13	41	0.1	91	38	A4780764
C	14	41	0.1	102	30	A4266556
C	15	41	0.1	103	81	B90619
C	16	40	0.1	96	22	R67088
C	17	40	0.1	105	28	A4078003
C	18	39	0.1	89	37	A4737634
C	19	39	0.1	100	26	W74144
C	20	39	0.1	102	34	AA481597
C	21	39	0.1	103	104	A0582186
C	22	39	0.1	105	81	B74378
C	23	38	0.1	74	28	A4078709
C	24	38	0.1	101	35	AA555145
C	25	38	0.1	104	36	AA614379
C	26	37	0.1	72	45	A1439123
C	27	37	0.1	95	34	AA457423
C	28	36	0.1	52	20	D25845
C	29	36	0.1	84	50	F34634
C	30	36	0.1	101	100	AQ280224
C	31	36	0.1	102	81	B80126
C	32	36	0.1	102	100	AQ319270
C	33	35	0.1	37	22	R70733
C	34	35	0.1	54	36	AA601314
C	35	35	0.1	84	34	AA501753
C	36	35	0.1	88	38	AA809831
C	37	35	0.1	93	21	R09732
C	38	35	0.1	94	23	R93021
C	39	35	0.1	95	24	H61099
C	40	35	0.1	95	35	AA578401
C	41	35	0.1	101	21	R17033
C	42	35	0.1	101	33	AA381369
C	43	35	0.1	101	34	AA494273
C	44	35	0.1	101	100	AQ278892
C	45	35	0.1	103	30	AA228795

ALIGNMENTS

RESULT 1
AA082835 59 bp mRNA EST 21-OCT-1996
LOCUS zn21g12.s1 Strataene neuroepithelium NT2RAMI 937234 Homo sapiens
DEFINITION cDNA clone IMAGE:348134 3' similar to contains Alu repetitive element; mRNA sequence.

ACCESSION
AA082835
VERSION
AA082835.1 GI:1624910
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 59)
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chippelli, B.,
Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevisan, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
JOURNAL
Genome Res. 6 (9), 807-828 (1996)
MEDLINE
97044478

COMMENT

On Sep 12, 1996 this sequence version replaced gi:1404580.
Contact: Willson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estel@wustl.edu
This clone is available royalty-free through LBNL; contact the
IMAGE Consortium (info@image.lbnl.gov) for further information.
Seq primer: -40M3 fwd. from Amersham.

FEATURES

Source
1. 59
/organism="Homo sapiens"
/db_xref="GDB:392650"
/db_xref="taxon:9606"
/clone="IMAGE:548134"
/clone_lib="Strataene neuroepithelium NT2RAMI 937234"
/dev_stage="NTera-2/RA+MI neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XbaI; Cloned unidirectionally. Primer: Oligo dT. NT2
(Ntera-2/cl.D1) precursor cells induced with Retinoid
Acid for 1 week, followed by 3 weeks in mitotic inhibitors
(Replate #2). Average insert size: 1.1 kb; Uni-ZAP XR
Vector: -5' adaptor sequence: 5' GAATTCGCGCAG 3' -3'
adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"

BASE COUNT

11 a 14 c 16 g 18 t

Query Match

Best Local Similarity 100.0%; Pred. No. 2.1e-08;

Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db

1 TTTAGTAGAGACGGGTTTACCATGTTAGCCAGAGTCTGATCTCTGACCTCGT 17491

RESULT 2
A0535244/c 103 bp DNA GSS 18-MAY-1999
LOCUS RPCI-11-317H22-TV RPCI-11 Homo sapiens genomic clone
DEFINITION RPCI-11-317H22, genomic survey sequence.
ACCESSION
A0535244
VERSION
A0535244.1 GI:4846934
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 103)
Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and
Venter, J.C.
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
Map Building
Unpublished (1997)
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbedet@igf.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pletier de Jong
(pletier@jg.med.bufileo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.bufileo.edu/ordering) or from
Research Genet cs (info@resgen.com). BAC end search page:
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: 17
Class: BAC ends.
Location/Qualifiers

FEATURES

Location/Qualifiers


```
source
1. 103
/organism="Homo sapiens"
/db_xref="GDB:7621533"
/db_xref="taxon:9606"
/clone_1lb="RPC1-11-317H22"
/clone_1lb="RPC1-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC11 Human Male BAC Library"

BASE COUNT      31 a      27 c      27 g      18 t

Query Match      0.2%; Score 59; DB 104; Length 103;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21703 AGAGACGGGTTTACCATGTTGGCCAGCGTGTCTGCAACTCTGACCTCAAGTATC 21761
|||||
Db 87 AGAGACGGGTTTACCATGTTGGCCAGCGTGTCTGCAACTCTGACCTCAAGTATC 29

RESULT 3
AI914923 72 bp mRNA EST 28-JUL-1999
LOCUS tr27e08.x1 NCI-CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2219558 3'
DEFINITION similar to contains Alu repetitive element; mRNA sequence.
ACCESSION AI914923
VERSION AI914923.1 GI:5634778
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 72)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On May 1, 1997 this sequence version replaced gi:2059395.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/db/ftp/image/image.html

JOURNAL
COMMENT
Unpublished (1997)
On May 1, 1997 this sequence version replaced gi:2059395.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/db/ftp/image/image.html

FEATURES
source
1. 72
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="NCI-CGAP_Ov23"
/tissue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/notes="Organ: ovary; Vector: pCMV-SPORE6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.35 kb. Tumor types include: mixed
Mullerian tumor, papillary serous, clear cell, spindle
cell. All are primary tumors, metastasis positive. Life
Technologies catalog #: 11534-013"

BASE COUNT      8 a      11 c      15 g      38 t

Query Match      0.2%; Score 52; DB 62; Length 72;

source
1. 102
/organism="Homo sapiens"
/db_xref="GDB:5491647"
/db_xref="taxon:9606"
/clone_1lb="669F15"
/clone_1lb="CIR-HSP"
/sex="Male"
/cell_type="Sperm"
/notes="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT      30 a      30 c      21 g      21 t

Query Match      0.2%; Score 48; DB 81; Length 102;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1313 TTTTGTATTATTAGTAGAGATGGGTTTCACCATGTTGCCAGGATG 1360
|||||
Db 59 TTTTGTATTATTAGTAGAGATGGGTTTCACCATGTTGCCAGGATG 12

RESULT 5
B48914/c 103 bp DNA GSS 08-APR-1999
LOCUS RPC111-4A12.TP RPC1-11 Homo sapiens genomic clone RPC1-11-4A12,
DEFINITION genomic survey sequence.
ACCESSION B48914
VERSION B48914.1 GI:2601151
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.

Best Local Similarity 100.0%; Pred. No. 3.1e-06;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21319 GCGCAGAGTGTGACTGTCTCAAAAAAAAAAAAAAAAAAAAAA 21370
|||||
Db 56 GCGCAGAGTGTGACTGTCTCAAAAAAAAAAAAAAAAAAAAAA 5

RESULT 4
B62983/c 102 bp DNA GSS 21-JUN-1998
LOCUS CIT-HSP-669F15.TP CIR-HSP Homo sapiens genomic clone 669F15,
DEFINITION genomic survey sequence.
ACCESSION B62983
VERSION B62983.1 GI:2636892
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 102)
Kim, D.-U., Adams, M.D. and Simon, M.I.
Determination of clone end sequences of human Bacterial Artificial
Chromosomes
Unpublished (1997)
Contact: Ung-Jin Kim
Caltech Genome Research Lab
California Institute of Technology
Division of Biology, MS 147-75, Pasadena, CA 91125, USA
Tel: 626 796 7066
Fax: 626 395 4901
Email: ung@ash.tree.caltech.edu
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.ligr.org/cdb/hungen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. 102
/organism="Homo sapiens"
/db_xref="GDB:5491647"
/db_xref="taxon:9606"
/clone_1lb="669F15"
/clone_1lb="CIR-HSP"
/sex="Male"
/cell_type="Sperm"
/notes="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT      30 a      30 c      21 g      21 t

Query Match      0.2%; Score 48; DB 81; Length 102;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1313 TTTTGTATTATTAGTAGAGATGGGTTTCACCATGTTGCCAGGATG 1360
|||||
Db 59 TTTTGTATTATTAGTAGAGATGGGTTTCACCATGTTGCCAGGATG 12

RESULT 5
B48914/c 103 bp DNA GSS 08-APR-1999
LOCUS RPC111-4A12.TP RPC1-11 Homo sapiens genomic clone RPC1-11-4A12,
DEFINITION genomic survey sequence.
ACCESSION B48914
VERSION B48914.1 GI:2601151
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
```

REFERENCE 1 (bases 1 to 103)
 AUTHORS Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C.
 TITLE Use of BAC End Sequences for Sequence-Ready Map Building
 JOURNAL Unpublished (1997)
 COMMENT Contact: Mark Adams
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: madams@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@jlong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 Source location/Qualifiers
 1..103
 /organism="Homo sapiens"
 /db_xref="GDB:7501163"
 /db_xref="taxon:9606"
 /clone="RPCI-11-4A12"
 /clone_1lb="RPCI-11"
 /sex="Male"
 /cell_type="Lymphocytes"
 /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"
 30 a 28 c 30 g 15 t

BASE COUNT
 30 a 28 c 30 g 15 t

ORIGIN

Query Match 0.2%; Score 48; DB 81; Length 103;
 Best Local Similarity 100.0%; Pred. No. 4.6e-05;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8728 ATCTCTGACCTCGATCGCCGCCCTCGCTCCCAAGAGCTGGG 8775
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 53 ATCTCTGACCTCGATCGCCGCCCTCGCTCCCAAGAGCTGGG 6

RESULT 6
 F29274/c 91 bp mRNA EST 13-MAY-1999
 LOCUS HSPD19063 HM3 Homo sapiens cDNA clone s4000074G04, mRNA sequence.
 DEFINITION F29274
 ACCESSION F29274
 VERSION F29274.1 GI:4814900
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 91)
 Lafranchi, G., Muraro, T., Caldara, F., Pacchioni, B., Pallavicini, A., Pandolfo, D., Toppo, S., Trevisan, S., Scarso, S. and Valle, G.
 Identification of 4370 expressed sequence tags from a 3'-end-specific cDNA library of human skeletal muscle by DNA sequencing and filter hybridization
 JOURNAL Genome Res. 6 (1), 35-42 (1996)
 MEDLINE 96276048
 COMMENT On Apr 7, 1998 this sequence version replaced gi:3034419.
 Contact: Valle G.
 CRIBI Biotechnology Centre
 University of Padua
 Via Trieste 75, 35121 Padua, Italy
 ABI Chromatograms and other information are available on WWW at
 http://grnp.bio.unipd.it.
 Location/Qualifiers
 1..91

FEATURES
 Source location/Qualifiers
 1..91

REFERENCE 1 (bases 1 to 105)
 AUTHORS Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C.
 TITLE Use of BAC End Sequences for Sequence-Ready Map Building
 JOURNAL Unpublished (1997)
 COMMENT Contact: Mark Adams
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: madams@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@jlong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 Source location/Qualifiers
 1..105
 /organism="Homo sapiens"
 /db_xref="GDB:487980"
 /db_xref="taxon:9606"
 /clone="s4000074G04"
 /clone_1lb="HM3"
 /sex="Female"
 /tissue_type="pectoral muscle (after mastectomy)"
 /note="Vector: pCDNAII (Invitrogen); Site_1: BstXI; Site_2: NotI; The library is not subtracted nor normalized. The first strand cDNA was primed with a biotinylated oligo-dT-NotI primer (5'-biotin-AACCCGCGCTGAGCGGCCCTTTT-3'). The ds cDNA was sonicated and size-selected in the range 350-550 bp. The 3' specific fragments were selected by streptavidin coated magnetic beads, ligated to non palindromic BstXI adapters. NotI digested and directionally cloned into BstXI-NotI cut pCDNAII vector."

BASE COUNT
 18 a 19 c 38 g 16 t

ORIGIN

Query Match 0.2%; Score 47; DB 50; Length 91;
 Best Local Similarity 100.0%; Pred. No. 0.0001;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 8621 TCCTGCTCAGCCTCCGAGTAGCTGAGGACACAGCCCTGCGACC 8667
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 63 TCCTGCTCAGCCTCCGAGTAGCTGAGGACACAGCCCTGCGACC 17

RESULT 7
 T94466 105 bp mRNA EST 24-MAR-1995
 LOCUS Y63502.1 Striagene lung (#937210) Homo sapiens cDNA clone IMAGE:119691 5' similar to contains Alu repetitive element; mRNA sequence.
 T94466
 VERSION T94466.1 GI:727954
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 105)
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapelli, B., Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Harkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierly-Meg, J., Trevaaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.
 Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (3), 807-828 (1996)
 MEDLINE 97044478
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 614
 High quality sequence stops: 70 Source: IMAGE Consortium, LNL This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 614 Std Error: 0.00
 Seq primer: M13RP1
 High quality sequence stop: 70.
 Location/Qualifiers
 1..105
 /organism="Homo sapiens"
 /db_xref="GDB:487980"
 /db_xref="taxon:9606"
 /clone="IMAGE:119691"

/clone_1lb="Stratagene lung (#937210)"
/sex="male"
/dev_stage="72 years"
/lab_host="SOLR cells (Kanamycin resistant)"
/note="Organ: lung; Vector: pBluescript SK-; Site.1:
EcoRI; Site.2: XhoI; Cloned unidirectionally. Primer:
0.11g0 dt. normal lung. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAGCAG
3' ~3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3' "

BASE COUNT 31 a 20 c 34 g 18 t 2 others

Query Match 0.2%; Score 47; DB 21; Length 105;
Best Local Similarity 100.0%; Pred. No. 9.4e-05;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23288 CCTCTGGGTTCAGTATCTCTGCTCAGCCTCTGAGTACGTG 25334
DB 88 CCTCTGGGTTCAGTATCTCTGCTCAGCCTCTGAGTACGTG 42

RESULT 8
LOCUS N84707 70 bp mRNA EST 01-APR-1996
DEFINITION J0579F Human fetal heart, Lambda ZAP Express Homo sapiens CDNA
clone J0579 5' similar to REPETITIVE ELEMENT ALU, mRNA sequence.
ACCESSION N84707
VERSION N84707.1 GI:1260332
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 70)
Liew.C.C.
CDNAs from fetal heart (1996)
Unpublished (1996)
Contact: Liew CC
Department of Laboratory Medicine and Pathobiology
University of Toronto
Banting Institute, 100 College St., Toronto, Ontario, M5G1L5
Tel: 4169788758
Fax: 4169785650
Email: liewcc@utcc.utoronto.ca
Seq primer: GAATTAACCTCTCCTAAAGG.
Location/Qualifiers

FEATURES
source 1..70
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="J0579"
/clone_1lb="human fetal heart, Lambda ZAP Express"
/lab_host="E. coli XL1-Blue"
/note="Vector: Lambda ZAP Express; Site.1: EcoRI; Site.2:
XhoI; mRNA was purified from human fetal hearts (8-10
weeks). CDNA was synthesized using a XhoI-Oligo dt
adaptor-primer. EcoRI adaptors were ligated, followed by
digestion with XhoI, for directional cloning into
predigested Lambda ZAP Express."

BASE COUNT 24 a 18 c 15 g 13 t

Query Match 0.2%; Score 44; DB 25; Length 70;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17445 GGGGTTTACCATGTTAGCCAGATGCTCGATCTCTGACACT 17488
DB 50 GGGGTTTACCATGTTAGCCAGATGCTCGATCTCTGACACT 7

RESULT 9

F24490/c
LOCUS F24490 80 bp mRNA EST 13-MAY-1999
DEFINITION HSPD10834 HM3 Homo sapiens CDNA clone s400003A06, mRNA sequence.
ACCESSION F24490
VERSION F24490.1 GI:4810116
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 80)
Lanfranchi,G., Muraro,T., Caldara,F., Pacchioni,B., Pallavicini,A.,
Pantolfo,D., Toppo,S., Trevisan,S., Scarso,S. and Valle,G.
Identification of 4370 expressed sequence tags from a
3'-end-specific cDNA library of human skeletal muscle by DNA
sequencing and filter hybridization
Genome Res. 6 (1), 35-42 (1996)

JOURNAL MEDLINE
COMMENT On Jun 5, 1998 this sequence version replaced gi:3188834.
96276048
Contact: Valle G.
CIRI Biotechnology Centre
University of Padua
Via Trieste 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
<http://grup.bio.unipd.it>.

FEATURES
source 1..80
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="s400003A06"
/clone_1lb="HM3"
/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/note="Vector: pCDNAII (Invitrogen); Site.1: BstXI;
Site.2: NotI. The library is not subtracted nor normalized.
Lanfranchi. This library is not subtracted nor normalized.
The first strand CDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-biotin-AACCGGCTCGAGCGGCGGCTTTTCTTTTCTTTT-3'). The
ds cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adaptors, NotI digested and
directionally cloned into BstXI-NotI cut pCDNAII vector."

BASE COUNT 18 a 21 c 19 g 22 t

Query Match 0.2%; Score 44; DB 50; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.00098;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19018 GGCCTCCCAAGTCTGGATTACGACGATGAGCCACTGCGCCC 19061
DB 80 GGCCTCCCAAGTCTGGATTACGACGATGAGCCACTGCGCCC 37

RESULT 10
LOCUS D20989 87 bp mRNA EST 30-JUL-1996
DEFINITION HUMGS01971 Human promyelocyte Homo sapiens CDNA clone mp0383 3',
RNA sequence.
ACCESSION D20989
VERSION D20989.1 GI:504809
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 87)
Okubo,K., Fukushima,A., Yoshii,J., Niiyama,T., Kojima,Y.,
Yoshinari,H., Arimoto,J. and Matsubara,K.
Gene expression of human promyelocytic cell line HL60 before and

REFERENCE 1 (bases 1 to 87)
AUTHORS Yoshinari,H., Arimoto,J. and Matsubara,K.
TITLE Gene expression of human promyelocytic cell line HL60 before and

JOURNAL
COMMENT

after induction of differentiation. A new application of 3' directed cDNA sequencing unpublished (1993)
Contact: Okubo, K., Fukushima, A., Yoshii, J., Niyama, T., Kojima, Y., Yoshinari, H., Arimoto, J. and Matsubara, K.
Institute for Molecular and Cellular Biology
Osaka University
3-1 Yamada-oka, Suita, Osaka 565, Japan.

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="mp0383"
/clone_lib="Human promyelocyte"
/note="Female, adult, cell_line = HL60, cell-type =
promyelocyte."
BASE COUNT      13 a      23 c      17 g      34 t
ORIGIN

```

Query Match	0.1%;	Score 43;	DB 20;	Length 87;
Best Local Similarity	100.0%;	Pred. No. 0.0019;		
Matches 43;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```
QY 25449 CCTACGCTCCCAAGTGTGGATTACAGGTGTAGCCACCA 25491
      |||
Db 11 CCTACGCTCCCAAGTGTGGATTACAGGTGTAGCCACCA 53
```

RESULT 11
AA835205

LOCUS	EST	DATE
AA835205	101 bp	23-FEB-1998
DEFINITION	ak64f01.s1 Barstedt pancreas HPLB1 Homo sapiens cDNA clone	
DESCRIPTION	IMAGE:1412689 3' similar to contains Alu repetitive element; contains element KBR repetitive element ; , mRNA sequence	

ACCESSION
VERSION

KEYWORDS	EST.
SOURCE	human.

ORGANISM

Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
AUTHORS

Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE
JOURNAL

On Nov 29, 1993 this sequence version replaced g1.b63d2b.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@lmbl.gov) for further information
Seq primer: -40m3 fwd. Et from Amersham.

FEATURES

source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image:1412689"
/clone_1lb="Barstead pancreas HPLRB1"
/sex="female"
/dev_stage="adult, 34 years"
/lab_host="DH10B"
/notes="Organ: pancreas; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; Site_1: EcoRI; Site_2: NotI;
1st strand cDNA was primed with a Not I - oligo(dT) primer"
15'
GGTACGACATCTGAAGTGGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT

```

```

3'1: double-stranded cDNA was ligated to Eco RI adaptors
[AAATCGACATCCTTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pTV73 vector.
Library constructed by Bob Barstead. "
BASE COUNT      14 a      36 c      27 g      24 t
ORIGIN

```

BASE COUNT
ORIGIN

Query Match: 0.1%; Score 43; DB 39; Length 101;
 Best Local Similarity: 100.0%; Pred. No. 0.0018;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8577 ATCTGGCTACTGCAAGCTCCGCTCCGGGTTCAAGCCATT 8613
 |||||
 Db 48 ATCTGGCTACTGCAAGCTCCGCTCCGGGTTCAAGCCATT 90

RESULT 12

LOCUS	AI864984	90 bp	MRNA	EST	30-AUG-1999
DEFINITION	WK06c11.x1	NCI_GCAP_Lym12	Homo sapiens	CDNA clone	IMAGE:2411540 3
DESCRIPTION	similar to contains Alu repetitive element; mRNA sequence.				

ACCESSION
VERSION
KEYWORDS

NETWORKS	ES1.
SOURCE	human.
ORGANISM	Homo sapiens

00000000000000000000

REFERENCE
1 (pages 1 to 90)
AUTHORS
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ccgap>

TITLE

TOPIC

On May 18, 1998 the

Contact: Robert Strausberg, Ph.D.
Tel.: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Life Technologies catalog #: 11347-015
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution Information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbir/image/image.html

FEATURES

1.90

```

/clone="IMAGE:2411340"
/clone_1lb="NCI-CGAP_Lym12"
/tissue_type="lymphoma, follicular mixed small and large
cell"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: pCMV-SPORT6; Site: 1;
Salt: Site 2: Not; Cloned unidirectionally. Primer:
Oligo dT. Average insert size 1.25 kb. Life Technologies
catalog #: 11547-015"
BASE COUNT      13 a      29 c      27 g      21 t
ORIGIN

```

Query Match	0.1%	Score 41	DB 61	Length 90
Best Local Similarity	100.0%	Pred. No.	0.0081	
Matches 41	Conservative 0	Mismatches 0	Indels 0	Gaps 0

```
QY      8577 ATCTGGCTCACTGCAAGCTCCGCCCTCCCGGGTTCACGCCA 8617  
       |||||  
Db      50 ATCTGGCTCACTGCAAGCTCCGCCCTCCCGGGTTCACGCCA 90
```

RESULT 13
AA780764

LOCUS AA780764 91 bp mRNA EST 05-FEB-1998
 DEFINITION AC6ff12.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone
 IMAGE:667791 3' similar to contains Alu repetitive element; mRNA
 sequence.
 ACCESSION AA780764
 VERSION AA780764.1 GI:2840095
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 91)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
 Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
 Martin, J., Moore, B., Schellenger, K., Steptoe, M., Tan, F.,
 Thelshing, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
 WashU-NCI human EST Project
 TITLE Unpublished (1997)
 JOURNAL On Sep 12, 1996 this sequence version replaced gi:1407381.
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: 40m13 fwd. ET from Amerisham.
 Location/Qualifiers
 1..91
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:667791"
 /clone_1lb="Stratagene fetal retina 937202"
 /sex="mixed"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Vector: pBluescript SK-; Site:1: EcoRI; Site:2:
 XhoI; Cloned unidirectionally. Primer: oligo dT. Pooled
 retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
 Vector: -5' adaptor sequence: 5' GAATTCGGCAGCG 3' -3'
 adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"

BASE COUNT 15 a 33 c 24 g 19 t
 ORIGIN

Query Match 0.1%; Score 41; DB 38; Length 91;
 Best Local Similarity 100.0%; Pred. No. 0.0081;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24778 CAATCGGCTCAGTCACACCTCCGCTCCGGGTCAACG 24818
 Db 43 CAATCGGCTCAGTCACACCTCCGCTCCGGGTCAACG 83

RESULT 14
 LOCUS AA226656 102 bp mRNA EST 21-AUG-1997
 DEFINITION nc19109.s1 NCI-CGAP_Prl Homo sapiens cDNA clone IMAGE:1008617
 similar to contains Alu repetitive element; mRNA sequence.
 ACCESSION AA226656
 VERSION AA226656.1 GI:1847964
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 102)
 NCI-CGAP http://www.nci.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1309483.
 Contact: Robert Strausberg, Ph.D.

Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,
 M.D., Michael Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: David B. Krizman, Ph.D.
 CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 282 Std Error: 0.00
 Seq primer: -41m13 fwd. ET from Amerisham
 High quality sequence stop: 82.
 Location/Qualifiers
 1..102
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1008617"
 /clone_1lb="NCI-CGAP_Prl"
 /sex="Male"
 /dev_stage="45 years old"
 /lab_host="DH10B"
 /note="Vector: PAMPI0; Site:1: NotI; Site:2: EcoRI; 1st
 strand cDNA was primed with oligo(dT)17 on 50 ng of
 DNase-treated, total cellular RNA obtained from
 5,000-10,000 microdissected, histologically normal
 prostate epithelial cells. Double-stranded cDNA was
 ligated to EcoRI adaptors, 5 cycles of PCR applied to the
 cDNA with an adaptor-specific primer, and the resulting
 PCR product subcloned into pAMP10 by the UDG-cloning
 method (Life Technologies). Average insert size is 600
 bp. NOTE: Not directionally cloned. This library was
 constructed by David Krizman."

BASE COUNT 13 a 32 c 26 g 30 t 1 others
 ORIGIN

Query Match 0.1%; Score 41; DB 30; Length 102;
 Best Local Similarity 100.0%; Pred. No. 0.0076;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18798 CTCTGTACCCAGCTGAGTGCAGTGCAGTCTCGCT 18838
 Db 21 CTCTGTACCCAGCTGAGTGCAGTGCAGTCTCGCT 61

RESULT 15
 LOCUS B90619/c 103 bp DNA GSS 25-JUN-1998
 DEFINITION CIT-HSP-2163G1.TR CIT-HSP Homo sapiens genomic clone 2163G1,
 genomic survey sequence.
 ACCESSION B90619
 VERSION B90619.1 GI:2973099
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 103)
 Adams, M.D., Rounsley, S.D., Zhao, S., Field, C.E., Bass, S., Linher, K.,
 Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,
 Simon, M. and Venter, J.C.
 Use of a random BAC End Sequence Database for Sequence-Ready Map
 Building (1998)
 TITLE Unpublished (1998)
 JOURNAL Other GSSs: CIT-HSP-2163G1.TF
 COMMENT Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208

Email: mdamams@ligr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page: http://www.ligr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source location/Qualifiers

1. 103
/organism="Homo sapiens"
/db_xref="GDB:7100363"
/db_xref="taxon:9606"
/clone="216361"
/clone_id="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pbelBAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT 27 a 31 c 27 g 18 t

ORIGIN

Query Match 0.18; Score 41; DB 81; Length 103;
Best Local Similarity 100.0%; Pred.No.0.0075;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17507 GCCTCCCAAGTGGGATTACAGGTGTGAGCCACGCGC 17547
|||||
Db 100 GCCTCCCAAGTGGGATTACAGGTGTGAGCCACGCGC 60

Search completed: April 1, 2000, 06:35:49
Job time: 271203 sec